

Docket No. 206272US2/btm

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Tsukasa KOHCHI

SERIAL NO: 09/838,311

FILED: April 20, 2001

FOR: METHOD AND APPARATUS FOR RECOGNIZING DOCUMENT IMAGE BY USE OF COLOR INFORMATION



GAU: 2621

EXAMINER:

2621  
#6  
DW  
11-29-01

REQUEST FOR PRIORITY

RECEIVED

ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

NOV 20 2001

Technology Center 2600

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number [US App No], filed [US App Dt], is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date of U.S. Provisional Application Serial Number, filed, is claimed pursuant to the provisions of 35 U.S.C. §119(e).
- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

COUNTRY

APPLICATION NUMBER

MONTH/DAY/YEAR

JAPAN

2000-124941

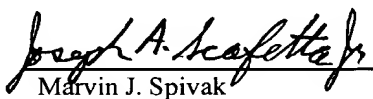
April 25, 2000

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number .  
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and  
(B) Application Serial No.(s)
  - ☐ are submitted herewith
  - ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.

  
Marvin J. Spivak

Registration No. 24,913

Joseph A. Scafetta, Jr.  
Registration No. 26,803



22850

Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 10/98)

PRIORITY DOCUMENT  
CERTIFIED COPY OF

09/838,311



日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出願年月日  
Date of Application:

2000年 4月25日

出願番号  
Application Number:

特願2000-124941

出願人  
Applicant(s):

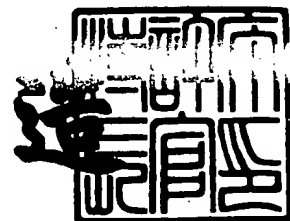
株式会社リコー

RECEIVED  
NOV 20 2001  
Technology Center 2600

2001年 4月20日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 0001155

【提出日】 平成12年 4月25日

【あて先】 特許庁長官 近藤 隆彦 殿

【国際特許分類】 G06T 7/00

【発明の名称】 文書画像認識方法、装置及びコンピュータ読み取り可能な記録媒体

【請求項の数】 32

【発明者】

【住所又は居所】 東京都大田区中馬込 1 丁目 3 番 6 号 株式会社リコー内

【氏名】 幸地 司

【特許出願人】

【識別番号】 000006747

【氏名又は名称】 株式会社リコー

【代理人】

【識別番号】 100070150

【弁理士】

【氏名又は名称】 伊東 忠彦

【先の出願に基づく優先権主張】

【出願番号】 特願2000- 32298

【出願日】 平成12年 2月 9日

【手数料の表示】

【予納台帳番号】 002989

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9911477

特 2 0 0 0 - 1 2 4 9 4 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文書画像認識方法、装置及びコンピュータ読み取り可能な記録媒体

【特許請求の範囲】

【請求項 1】 文書画像を認識する文書画像認識方法であって、  
文書画像をデジタル画像として入力する入力ステップと、  
該文書画像の背景色を特定する背景色特定ステップと、  
該背景色を用いて該文書画像から背景領域以外の画素を抽出する抽出ステップと、  
該画素を統合して連結成分を生成する生成ステップと、  
該連結成分を少なくとも形状特徴を用いて所定の領域に分類して、該文書画像の領域識別の結果を得る分類ステップと  
を有することを特徴とする文書画像認識方法。

【請求項 2】 前記領域識別された文書画像を 2 値化し、2 値画像を生成するステップと、  
該 2 値画像の領域を所定の領域に分類し、その結果を前記領域識別の結果と照合して該領域識別の結果を補正する補正ステップと、  
文字領域を文字認識する認識ステップと  
を更に有する請求項 1 に記載の文書画像認識方法。

【請求項 3】 前記背景色特定ステップは、  
前記文書画像の色のクラスタリングを行うクラスタリングステップと、  
クラスタリングにより得られた最大クラスタの代表色を該文書画像の背景色とするステップと  
を有する請求項 1 に記載の文書画像認識方法。

【請求項 4】 前記クラスタリングステップは、  
ある間隔を空けて画素をサンプリングするステップと、  
該画素の近傍を平滑化した画素値を用いて前記クラスタリングを行うステップと  
を有する請求項 3 に記載の文書画像認識方法。

【請求項 5】 前記文書画像認識方法は更に前記文書画像を縮小する縮小ステップを有し、該縮小ステップは、

文書画像を複数のブロックに分割するステップと、

ブロック内の代表色を求めるステップと、

該代表色と前記背景色とを比較して、該ブロックの縮小後の色を決定し、該ブロックを該色の画素に縮小するステップと

を有する請求項 1 に記載の文書画像認識方法。

【請求項 6】 前記ブロックは、3 画素×3 画素又は 4 画素×4 画素の格子である請求項 5 に記載の文書画像認識方法。

【請求項 7】 前記抽出ステップは、

前記背景色と注目画素の各色値の差が所定の値より大きい場合に該注目画素は背景領域以外の画素であると判断するステップを有する請求項 1 に記載の文書画像認識方法。

【請求項 8】 前記領域識別によって識別された図又は写真の矩形領域を特定色で塗り潰した文書画像を生成するステップと、

該文書画像に 2 値化処理を施して得られた 2 値画像に対して文字認識を行うステップと

を更に有する請求項 1 に記載の文書画像認識方法。

【請求項 9】 前記分類ステップにおいて分類された所定の矩形領域に対して再帰的に該分類ステップの処理を実行する請求項 1 に記載の文書画像認識方法。

【請求項 10】 文書画像を認識する文書画像認識方法であって、

文書画像をデジタル画像として入力するステップと、

該文書画像をカラー領域識別するステップと、

カラー領域識別された領域毎の二値画像を生成するステップと、

領域毎の二値画像を統合して一枚の二値画像を生成し、該二値画像の領域識別を行うステップと、

該二値画像の領域識別結果と前記カラー領域識別結果とを照合し、必要に応じてフィードバック処理を所定の条件を満たすまで又は所定回数行って二値画像と

領域識別結果を得るステップと

を有することを特徴とする文書画像認識方法。

【請求項 1 1】 カラー領域識別結果と二値画像の領域識別結果との照合の結果、文書のある範囲にて所定の条件を満たさない場合に前記フィードバック処理が行われ、該フィードバック処理は、

該範囲を包含する領域を作成し、該領域に対して再度カラー領域識別、二値化及び二値領域識別を行い、両者の領域識別結果を照合するステップを有する請求項 1 0 に記載の文書画像認識方法。

【請求項 1 2】 カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にて一方の領域識別で文字行が抽出され、もう一方の領域識別結果では文字矩形が得られなかった場合に、前記フィードバック処理が行われ、該フィードバック処理は、

該文字矩形の文字色を特定し、文字色にばらつきがなければ、該範囲には文字が存在すると判定し、特定した文字色を用いて再度カラー領域識別、二値化及び二値領域識別を行い、両者の領域識別結果を照合するステップを有する請求項 1 0 に記載の文書画像認識方法。

【請求項 1 3】 前記フィードバック処理は、

カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にてカラー領域識別では抽出された文字行が二値領域識別の結果では存在しない場合に、該当する文字行を包含する領域を作成し、二値化及び二値領域識別を行い、該領域識別結果とカラー領域識別結果とを照合するステップを含む請求項 1 2 に記載の文書画像認識方法。

【請求項 1 4】 カラー領域識別結果と二値画像の領域識別結果との照合の結果、双方の領域識別結果で所定行数以上連続してレイアウト特徴が異なる文字行が存在する場合に、前記フィードバック処理が行われ、該フィードバック処理は、

該当する文字行を包含する領域を作成し、二値化及び二値領域識別し、該領域識別結果とカラー領域識別結果とを照合するステップを有する請求項 1 0 に記載の文書画像認識方法。

【請求項 1 5】 前記カラー領域識別された領域毎の二値画像を生成するステップにおいて、

文字部には画像分割型の二値化方式を適用し、罫線部、図、写真領域には判別分析型の二値化方式を適用する請求項 1 0 に記載の文書画像認識方法。

【請求項 1 6】 前記カラー領域識別は、請求項 1 に記載の文書画像認識方法により行う請求項 1 0 ないし 1 5 のうちいずれか 1 項に記載の文書画像認識方法。

【請求項 1 7】 文書画像を認識する文書画像認識装置であって、  
文書画像をデジタル画像として入力する入力手段と、  
該文書画像の背景色を特定する背景色特定手段と、  
該背景色を用いて該文書画像から背景領域以外の画素を抽出する抽出手段と、  
該画素を統合して連結成分を生成する生成手段と、  
該連結成分を少なくとも形状特徴を用いて所定の領域に分類して、該文書画像の領域識別の結果を得る分類手段と  
を有することを特徴とする文書画像認識装置。

【請求項 1 8】 前記領域識別された文書画像を 2 値化し、2 値画像を生成する手段と、

該 2 値画像の領域を所定の領域に分類し、その結果を前記領域識別の結果と照合して該領域識別の結果を補正する補正手段と、  
文字領域を文字認識する認識手段と  
を更に有する請求項 1 7 に記載の文書画像認識装置。

【請求項 1 9】 前記背景色特定手段は、  
前記文書画像の色のクラスタリングを行うクラスタリング手段と、  
クラスタリングにより得られた最大クラスタの代表色を該文書画像の背景色とする手段と

を有する請求項 1 7 に記載の文書画像認識装置。

【請求項 2 0】 前記クラスタリング手段は、  
ある間隔を空けて画素をサンプリングする手段と、  
該画素の近傍を平滑化した画素値を用いて前記クラスタリングを行う手段と



を有する請求項 1 9 に記載の文書画像認識装置。

【請求項 2 1】 前記文書画像認識装置は更に前記文書画像を縮小する縮小手段を有し、該縮小手段は、

文書画像を複数のブロックに分割する手段と、

ブロック内の代表色を求める手段と、

該代表色と前記背景色とを比較して、該ブロックの縮小後の色を決定し、該ブロックを該色の画素に縮小する手段と

を有する請求項 1 7 に記載の文書画像認識装置。

【請求項 2 2】 前記ブロックは、3 画素×3 画素又は 4 画素×4 画素の格子である請求項 2 1 に記載の文書画像認識装置。

【請求項 2 3】 前記抽出手段は、

前記背景色と注目画素の各色値の差が所定の値より大きい場合に該注目画素は背景領域以外の画素であると判断する手段を有する請求項 1 7 に記載の文書画像認識装置。

【請求項 2 4】 前記領域識別によって識別された図又は写真の矩形領域を特定色で塗り潰した文書画像を生成する手段と、

該文書画像に 2 値化処理を施して得られた 2 値画像に対して文字認識を行う手段と

を更に有する請求項 1 7 に記載の文書画像認識装置。

【請求項 2 5】 前記分類手段において分類された所定の矩形領域に対して再帰的に該分類手段の処理を実行する請求項 1 7 に記載の文書画像認識装置。

【請求項 2 6】 文書画像を認識する文書画像認識装置であって、

文書画像をデジタル画像として入力する手段と、

該文書画像をカラー領域識別する手段と、

カラー領域識別された領域毎の二値画像を生成する手段と、

領域毎の二値画像を統合して一枚の二値画像を生成し、該二値画像の領域識別を行う手段と、

該二値画像の領域識別結果と前記カラー領域識別結果とを照合し、必要に応じてフィードバック処理を所定の条件を満たすまで又は所定回数行って二値画像と

領域識別結果を得る手段と

を有することを特徴とする文書画像認識装置。

【請求項 2 7】 カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にて一方の領域識別で文字行が抽出され、もう一方の領域識別結果では文字矩形が得られなかった場合に、前記フィードバック処理が行われ、該フィードバック処理において、

該文字矩形の文字色を特定し、文字色にばらつきがなければ、該範囲には文字が存在すると判定し、特定した文字色を用いて再度カラー領域識別、二値化及び二値領域識別を行い、両者の領域識別結果を照合する請求項 2 6 に記載の文書画像認識装置。

【請求項 2 8】 前記フィードバック処理は、

カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にてカラー領域識別では抽出された文字行が二値領域識別の結果では存在しない場合に、該当する文字行を包含する領域を作成し、二値化及び二値領域識別を行い、該領域識別結果とカラー領域識別結果とを照合する処理を含む請求項 2 7 に記載の文書画像認識装置。

【請求項 2 9】 カラー領域識別結果と二値画像の領域識別結果との照合の結果、双方の領域識別結果で所定行数以上連続してレイアウト特徴が異なる文字行が存在する場合に、前記フィードバック処理が行われ、該フィードバック処理において、

該当する文字行を包含する領域を作成し、二値化及び二値領域識別し、該領域識別結果とカラー領域識別結果とを照合する請求項 2 6 に記載の文書画像認識装置。

【請求項 3 0】 コンピュータに文書画像認識処理を実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体であって、

文書画像をデジタル画像として入力する入力手順と、

該文書画像の背景色を特定する背景色特定手順と、

該背景色を用いて該文書画像から背景領域以外の画素を抽出する抽出手順と、

該画素を統合して連結成分を生成する生成手順と、

該連結成分を少なくとも形状特徴を用いて所定の領域に分類して、該文書画像の領域識別の結果を得る分類手順と

をコンピュータに実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 3 1】 前記領域識別された文書画像を 2 値化し、2 値画像を生成する手順と、

該 2 値画像の領域を所定の領域に分類し、その結果を前記領域識別の結果と照合して該領域識別の結果を補正する補正手順と、

文字領域を文字認識する認識手順と

を更にコンピュータに実行させる請求項 3 0 に記載のプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項 3 2】 コンピュータに文書画像認識処理を実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体であって、

文書画像をデジタル画像として入力する手順と、

該文書画像をカラー領域識別する手順と、

カラー領域識別された領域毎の二値画像を生成する手順と、

領域毎の二値画像を統合して一枚の二値画像を生成し、該二値画像の領域識別を行う手順と、

該二値画像の領域識別結果と前記カラー領域識別結果とを照合し、必要に応じてフィードバック処理を所定の条件を満たすまで又は所定回数行って二値画像と領域識別結果を得る手順と

をコンピュータに実行させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、文書画像の領域識別技術に関する。特に、カラー文書画像中の写真等の領域を分離して、文字や罫線等を識別し、文字を認識し、レイアウトや色情報を含めた文書の再現に役立てるための技術に関する。

## 【 0 0 0 2 】

## 【従来の技術】

OCR処理において文字を認識するためには、その前段階として正確に領域を識別する必要がある。カラー原稿に対するOCR処理等のための領域識別においては、画像を2値化して処理を行うことが一般的であり、精度良く領域識別することはできなかった。このように、従来の文字認識の分野においては、2値化処理を用いる場合が多く、色情報を積極的に用いることは少ない。例えば、特開平07-998581「画像処理装置」（オリンパス）に開示されているように、カラー画像が入力された場合には、まず最初に2値化して、生成された2値画像に対して領域識別や文字認識処理がなされる。

## 【 0 0 0 3 】

また、カラー文書画像の領域を識別する技術に関しては次に示すような従来技術がある。

## 【 0 0 0 4 】

特開平05-64016「画像圧縮装置」（キャノン）には、入力カラー文書画像を効率よく圧縮するために、自然画（写真）と線画（文字、イラスト、グラフ）を分離する装置が開示されている。しかし、特開平05-64016では、文字領域と文字以外の領域とを効率的に分離する技術は示されていない。従って、この技術による領域識別結果をカラー文書のOCRにそのまま適応することは困難である。

## 【 0 0 0 5 】

また、特開平06-339019「離散的コサイン変換による文書画像の領域分離方式」（NTT）には、画像を周波数成分で表し、離散的コサイン変換を用いて文字と写真領域の分離を行う技術が開示されている。ここで、離散的コサイン変換係数の中から文字と写真を特徴づける係数を精度よく抽出するためには、高解像度かつ高品質のカラー画像が必要である。従って、特開平06-339019によると、巨大なディスク・メモリー容量が消費されてしまうと同時にかかりの実行時間がかかってしまう恐れがある。

## 【 0 0 0 6 】

また、特開平 7-168939「カラー画像領域判定装置」（松下ソフトリサーチ）には、色相のヒストグラムを利用して、文書画像中のカラー写真領域と色の少ない図形または文字領域を判定する技術が開示されている。しかし、特開平 7-168939においては、領域判定にあたり色相に着目していることから、新聞などの白黒・グレー原稿には全く無力である。

## 【0007】

さて、OCRに関して現在多くの手法が提案されているが、そのほとんどが二値画像を対象としたものである。従って二値画像の状態が認識精度を大きく左右する。一方、カラー化された文書では、単に使用される色が増えるだけではなく文書のレイアウト構造も複雑になる傾向がある。しかしながら、カラー画像から生成された二値画像に対して従来の方法で領域識別処理を行ったとしても、実際には十分な精度が得られない場合が多い。

## 【0008】

文字認識に好適な二値画像を生成するために数多くの手法が提案されており、その中の代表的な技術として判別分析法である。

## 【0009】

また、特開平 10-143608は濃淡画像を適当な閾値で二値化し、できた二値画像上の平均線幅を計算してその値が規定範囲外にある場合には、文字認識に不適であると判断して二値化をやり直すことにより、OCRに好適な二値画像を生成する。

## 【0010】

## 【発明が解決しようとする課題】

上述の通り、従来の技術では、カラー原稿を文字認識して、レイアウトや色情報を含めて精度良く再現することができない。

## 【0011】

また、カラー化された文書に対してはカラー特有の問題が混在する場合が多く、OCRに好適な二値画像を精度良く生成することは困難であった。

## 【0012】

本発明は上記の点に鑑みてなされたものであり、カラー文書画像及び白黒・グ

レー画像の領域を精度良く、効率的に識別する文書画像認識方法、装置及びコンピュータ読み取り可能な記録媒体を提供することを目的とする。

【 0 0 1 3 】

更に、カラー特有の問題が混在する場合であってもOCRに好適な画像を精度良く生成する文書画像認識方法、装置及びコンピュータ読み取り可能な記録媒体を提供することを目的とする。

【 0 0 1 4 】

【課題を解決するための手段】

上記の課題を解決するために、本発明は次のように構成される。

【 0 0 1 5 】

請求項 1 に記載の発明は、文書画像を認識する文書画像認識方法であって、文書画像をデジタル画像として入力する入力ステップと、該文書画像の背景色を特定する背景色特定ステップと、該背景色を用いて該文書画像から背景領域以外の画素を抽出する抽出ステップと、該画素を統合して連結成分を生成する生成ステップと、該連結成分を少なくとも形状特徴を用いて所定の領域に分類して、該文書画像の領域識別の結果を得る分類ステップとを有する。

【 0 0 1 6 】

本発明によれば、背景色を特定し、その背景色を用いてボトムアップ的に文書画像を領域識別結果を得るので、色情報を用いない従来技術と比較して、精度良くカラー画像及び白黒・グレー画像の領域識別を行うことができる。

【 0 0 1 7 】

請求項 2 に記載の発明は、請求項 1 の記載において、前記領域識別された文書画像を 2 値化し、2 値画像を生成するステップと、該 2 値画像の領域を所定の領域に分類し、その結果を前記領域識別の結果と照合して該領域識別の結果を補正する補正ステップと、文字領域を文字認識する認識ステップとを更に有する。

【 0 0 1 8 】

本発明によれば、精度良く図や写真領域が文字領域と区別されるので、従来より精度良く文字認識を行うことができる。

【 0 0 1 9 】

請求項 3 に記載の発明は、請求項 1 の記載において、前記背景色特定ステップは、前記文書画像の色のクラスタリングを行うクラスタリングステップと、クラスタリングにより得られた最大クラスタの代表色を該文書画像の背景色とするステップとを有する。

【 0 0 2 0 】

本発明によれば、最大クラスタの代表色を背景色とすることで、カラー画像の背景色を効率良く求めることができ、背景色を用いた領域識別の処理を行うことが可能となる。

【 0 0 2 1 】

請求項 4 に記載の発明は、請求項 3 の記載において、前記クラスタリングステップは、ある間隔を空けて画素をサンプリングするステップと、該画素の近傍を平滑化した画素値を用いて前記クラスタリングを行うステップとを有する。

【 0 0 2 2 】

本発明によれば、ある間隔を空けて画素をサンプリングするので、処理の高速化を図ることができる。

【 0 0 2 3 】

請求項 5 に記載の発明は、請求項 1 の記載において、更に前記文書画像を縮小する縮小ステップを有し、該縮小ステップは、文書画像を複数のブロックに分割するステップと、ブロック内の代表色を求めるステップと、該代表色と前記背景色とを比較して、該ブロックの縮小後の色を決定し、該ブロックを該色の画素に縮小するステップとを有する。

【 0 0 2 4 】

本発明によれば、背景色を用いることによって、カラー画像の場合でも領域識別にふさわしい縮小画像を求めることが可能となる。また、現画像を縮小することで、画像へのアクセスコストを大幅に削減することが可能になる。

【 0 0 2 5 】

請求項 6 に記載の発明は、請求項 5 の記載において、前記ブロックは、3 画素×3 画素の格子であるとする。3 画素×3 画素又は 4 画素×4 画素とすることによって、一般的なカラーキャナで入力された文書画像に対して、適切な領域識

別を行うことが可能な縮小画像を求めることができる。また、格子状とすることで実装が容易になる。

【 0 0 2 6 】

請求項 7 に記載の発明は、請求項 1 の記載において、前記抽出ステップは、前記背景色と注目画素の各色値の差が所定の値より大きい場合に該注目画素は背景領域以外の画素であると判断するステップを有する。

【 0 0 2 7 】

このように比較的簡単な演算によって背景と文書要素を分離することが可能となるので、処理速度の向上を図れる。

【 0 0 2 8 】

請求項 8 に記載の発明は、請求項 1 の記載において、前記領域識別によって識別された図又は写真の矩形領域を特定色で塗り潰した文書画像を生成するステップと、該文書画像に 2 値化処理を施して得られた 2 値画像に対して文字認識を行うステップとを更に有する。

【 0 0 2 9 】

本発明によれば、図・写真等の領域の周辺にある文字部分を精度良く抽出することができる。

【 0 0 3 0 】

請求項 9 に記載の発明は、請求項 1 の記載において、前記分類ステップにおいて分類された所定の矩形領域に対して再帰的に該分類ステップの処理を実行する。

【 0 0 3 1 】

本発明によれば、文書全体の背景色と異なる背景色を有する部分領域内も正確に領域識別される。

【 0 0 3 2 】

請求項 1 0 に記載の発明は、文書画像を認識する文書画像認識方法であって、文書画像をデジタル画像として入力するステップと、該文書画像をカラー領域識別するステップと、カラー領域識別された領域毎の二値画像を生成するステップと、領域毎の二値画像を統合して一枚の二値画像を生成し、該二値画像の領域識



別を行うステップと、該二値画像の領域識別結果と前記カラー領域識別結果とを照合し、必要に応じてフィードバック処理を所定の条件を満たすまで又は所定回数行って二値画像と領域識別結果を得るステップとを有する。

## 【 0 0 3 3 】

本発明によれば、OCR前処理（領域識別、二値化）へのフィードバック機構を有することとしたので、従来のカラー文書認識技術では難しかったカラー文書特有の問題に対しても対応可能となる。例えば、従来技術では難しかった各セルが色分けされた表や、グラデーションが大きく変化する地肌に書かれた文字、あるいはこれらが混在する場合でも精度よくOCRすることができる。

## 【 0 0 3 4 】

請求項 1 1 に記載の発明は、請求項 1 0 の記載において、カラー領域識別結果と二値画像の領域識別結果との照合の結果、文書のある範囲にて所定の条件を満たさない場合に前記フィードバック処理が行われ、該フィードバック処理は、

該範囲を包含する領域を作成し、該領域に対して再度カラー領域識別、二値化及び二値領域識別を行い、両者の領域識別結果を照合するステップを有する。

## 【 0 0 3 5 】

本発明によっても、請求項 1 0 と同様の効果を得ることができる。

## 【 0 0 3 6 】

請求項 1 2 に記載の発明は、請求項 1 0 の記載において、カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にて一方の領域識別で文字行が抽出され、もう一方の領域識別結果では文字矩形が得られなかった場合に、前記フィードバック処理が行われ、該フィードバック処理は、該文字矩形の文字色を特定し、文字色にばらつきがなければ、該範囲には文字が存在すると判定し、特定した文字色を用いて再度カラー領域識別、二値化及び二値領域識別を行い、両者の領域識別結果を照合するステップを有する。

## 【 0 0 3 7 】

本発明によれば、領域単位ではなく文字行で領域識別結果を比較し、文字行が両方にあるかどうかを調べるので、従来避けられなかった行から領域への統合誤りを回避することができる。その際、文字色を特定して文字色のばらつきを調べ

ることで、比較対象範囲に文字が存在するか否かを精度よく判定でき、その結果文字が存在する範囲だけをフィードバックすることで周囲への悪影響がなく文字抽出精度が向上する。

## 【 0 0 3 8 】

請求項 1 3 に記載の発明は、請求項 1 2 の記載において、前記フィードバック処理は、カラー領域識別結果と二値画像の領域識別結果との照合の結果、ある範囲にてカラー領域識別では抽出された文字行が二値領域識別の結果では存在しない場合に、該当する文字行を包含する領域を作成し、二値化及び二値領域識別を行い、該領域識別結果とカラー領域識別結果とを照合するステップを含む。

## 【 0 0 3 9 】

本発明によれば、領域単位ではなく文字行単位で領域識別結果を比較することにより、従来避けられなかった行から領域への統合誤りを回避することができる。

その際、カラー情報を積極的に用いるカラー領域識別での文字抽出結果を優先的に採用することで、二値化が原因で従来 O C R でできなかった文字をフィードバックして抽出する。

## 【 0 0 4 0 】

請求項 1 4 に記載の発明は、請求項 1 0 の記載において、カラー領域識別結果と二値画像の領域識別結果との照合の結果、双方の領域識別結果で所定行数以上連続してレイアウト特徴が異なる文字行が存在する場合に、前記フィードバック処理が行われ、該フィードバック処理は、該当する文字行を包含する領域を作成し、二値化及び二値領域識別し、該領域識別結果とカラー領域識別結果とを照合するステップを有する。

## 【 0 0 4 1 】

本発明によれば、領域単位ではなく文字行単位で領域識別結果を比較することにより、従来避けられなかった行から領域への統合誤りを回避することができる。

その際、対象範囲の複数行についてそれぞれレイアウト特徴を比較することで、どちらかの領域識別結果での誤りを正確に検出することができ、該当範囲だけを

フィードバックすることで周囲への悪影響なく文字抽出精度が向上する。

【 0 0 4 2 】

請求項 1 5 に記載の発明は、請求項 1 0 の記載において、前記カラー領域識別された領域毎の二値画像を生成するステップにおいて、文字部には画像分割型の二値化方式を適用し、罫線部、図、写真領域には判別分析型の二値化方式を適用する。

【 0 0 4 3 】

画像分割型二値化方法は、悪条件下（グラデーション、周辺ノイズなど）でも、人間が認識可能なレベルで文字と背景を分離することが可能であるので、文字領域の二値化に適している。また、判別分析法を用いて罫線、図・写真領域を二値化することで、特に二値化の閾値を多少濃い目に設定することにより罫線がかすれず図や写真も一塊になり、領域識別にとって有利な二値画像となる

請求項 1 6 に記載の発明は、請求項 1 0 ないし 1 5 のうちいずれか 1 項の記載において、前記カラー領域識別は、請求項 1 に記載の文書画像認識方法により行うこととする。これにより、精度良くカラー領域識別できるとともに、文字色を利用しやすくなる。

【 0 0 4 4 】

請求項 1 7 ～ 2 9 に記載された発明は、本発明の文書画像認識方法の実施に適した文書画像認識装置であり、上述した作用及び効果と同様の作用効果を得ることができる。

【 0 0 4 5 】

請求項 3 0 ～ 3 2 に記載された発明は、本発明の文書画像認識方法をコンピュータ上で実行するためのプログラムを記録したコンピュータ読み取り可能な記録媒体であり、この発明によっても、上述した作用及び効果と同様の作用効果を得ることができる。

【 0 0 4 6 】

上述の通り、本発明によれば、ディスクやメモリーの過剰な消費を押さえて、かつ高速な画像認識装置及び方法を提供することができる。

【 0 0 4 7 】

また多彩な色が用いられたカラー文書だけではなく、新聞等のような白黒・グレイ画像に対しても精度良く領域識別を行うことができる。

【0048】

更に、カラー文書特有の問題を有するカラー文書に対しても精度よくOCRすることができる。

【0049】

【発明の実施の形態】

(第1の実施例)

以下、図を参照して本発明における第1の実施例を説明する。図1は本発明の文書画像認識装置の構成図である。同図に示すように、本発明の文書画像認識装置は、CPU（中央処理装置）100、メモリ101、スキャナ102、通信装置103、表示装置104、ハードディスク105、キーボード106、CD-ROMドライブ107を有する。ハードディスク105は文書画像認識装置の外部装置として接続してもよいし、内部装置として有していてもよい。CPU100は文書画像認識装置の全体を制御する。メモリ101はCPU100で処理するデータやプログラムを保持する。通信装置103は文書画像認識装置をLAN等のネットワークに接続するための制御を行う。キーボード106はデータを入力する装置である。CD-ROMドライブ107はCD-ROM等を駆動し、読み書きを行う。スキャナ102は文書画像認識装置の外部装置として接続され、文書をカラーデジタル画像として入力する機能を有する。CD-ROMドライブ107の他、MOドライブ、FDドライブ、DVDドライブ等でもよい。

【0050】

本発明において後述する処理を実行するプログラムは、例えばCD-ROMに格納され、CD-ROMドライブ107を介してハードディスク105にロードされる。プログラムが起動されると、所定のプログラムがメモリ101に展開され、処理が実行される。動作の概要は次の通りである。

【0051】

まず、スキャナ102を介して紙文書の内容がカラーデジタル画像として入力され、メモリ101又はハードディスク105に格納される。次に、以下で詳述

する領域識別がCPU100により行われ、領域識別結果に対して、例えばOCR処理が行われる。

#### 【0052】

図2は本発明の第1の実施例における処理の全体を示すフローチャートである。図2を用いて全体の処理を説明し、その後各処理について詳細に説明する。

まず、ステップS101として、紙文書であるカラー文書画像を入力する。この入力、は、スキャナーで、あるいはネットワークを介してカラーデジタル画像として入力することが可能である。本実施例では色数を24bitフルカラーと仮定するが、本発明は、基本的に画像の色数や解像度には依存しない。また、本実施例では、扱う画像はビットマップ形式で表現されていると仮定している。従って、画像を周波数成分で表現するJPEGのような画像が入力された場合には、最初に画像全体をビットマップ形式に変換する。本実施例における画像形式としては、例えば、200dpi、24bitカラー、WindowsBMP形式、非圧縮、色表現RGB座標系とすることができる。

#### 【0053】

ステップS102として、文書の背景色を検出する。次に、ステップS103にて画像へのアクセスコストを押さえるために原画像を縮小する。ステップS104において、縮小画像に対して背景色と異なる画素からランを生成し、次に、ランを統合して矩形を生成して（ステップS105）、最後に小さな矩形を統合する（ステップS106）。なお、背景色特定の処理（ステップS102）と縮小画像生成の処理（ステップS103）は逆でもよい。

#### 【0054】

次に、ステップS107として、生成された矩形を写真／テキスト／罫線などに分類し、ステップS108において、前記分類された矩形の座標値を、原画像上における実際の値にあわせるために適宜拡大変換する。上記のステップS101～ステップS108の処理がカラー文書画像の領域識別のための処理である。

本発明では、更に原画像に戻り、ステップS109で、前記抽出した写真矩形を黒、あるいは指定された色で塗り潰したカラー画像を生成する。次に、従来の文字認識技術を適応するために、ステップS110にて、前記生成したカラー画

像を変換して2値400dpi相当の画像を生成する。画像の2値化には種々の方法が適用可能であり、例えば、大津の方法（判別分析法）等を適用することができる。

## 【0055】

次に、ステップS111において、生成した2値400dpi相当の画像に対して、従来の文字認識手法でOCR処理する。必要に応じて、OCR処理の前に従来の2値画像に対する領域識別を施して、前記カラー文書画像から直接得られた領域識別結果とを照合することにより文書の領域識別結果を補正してもよい。特に、図・写真領域の識別に関しては、本発明における色情報を用いた領域識別結果を採用すればよい。

## 【0056】

最後に、ステップS112として、OCR処理の結果とカラー画像から直接分離した写真領域を組み合わせて、例えばHTML形式で文書を再現する。すなわち、OCR結果から文字コードが、図・写真等の領域識別結果からはいずれもオリジナル画像に対する座標値が分かるので、文書レイアウトを再現することは容易である。例えば、PDF、PS、RTF形式等で文書再現を行うことができ、その際、写真・図の矩形はオリジナルのカラーで表現する。また、読み順を考慮して、HTML形式等で再現すれば、論理的に文書を再現でき、なおかつ、写真・図はオリジナルのカラーで表現できるので、Web上の利用効果が大きい。

## 【0057】

次に、与えられた矩形領域（画像全体も可）の背景色を特定するための背景色特定処理（上記のステップS102に相当する）について詳細に説明する。図3は、背景色特定処理を示すフローチャートである。

## 【0058】

ステップS201として、処理対象となる矩形（Aとする）を入力する。与えられた矩形が原画像全体ならば、ステップS202において、スキャン時に混入が予想される周辺ノイズを避けるために、周辺ノイズ除去を行う。例えば、対象とする範囲をすこし狭めてもよい。

## 【0059】

次に、矩形Aにおける色の分布具合を調べるために、矩形Aで用いられている色のクラスタリングを行う（ステップS203～ステップS209）。

【0060】

まず、ステップS203にて、例えば、64画素ごとにサンプリングする。本実施例では、処理の高速化を図るために、矩形A全体を走査するのではなく一定の間隔をあけて画素をサンプリングする。次に、サンプリングされた画素P（注目画素という）からクラスタリングに用いる画素値を求める。ここでは文書画像の文字部のエッジを保持するために、Pの近傍を参照して平滑化された画素値を得る。本実施例では注目画素を中心とした5×5ブロックを近傍としている（ステップS204）。

【0061】

注目画素Pを中心とした近傍5×5ブロック内を参照した平滑化は、例えば「コンピュータ画像処理入門」田村秀行著、総研出版、P. 105図4. 12に記載された方法で行う（ステップS205～ステップS207）。すなわち、ステップS205において、注目画素Pに対してブロック内の9種類のエリアを定義して、それぞれのエリア内での濃度の分散を求める。「コンピュータ画像処理入門」における9種類のエリアの例を図4に示す。

【0062】

続いて、ステップS206において、分散が最小となる近傍エリアを選択して、そのエリア内での代表色を各画素値の平均値等を用いて決定する（ステップS207）。

【0063】

次に、画像領域全体を走査したかどうかを調べ（ステップS208）、まだ全体を走査していない場には、再びサンプリング（ステップS203）から処理を行う。全体を走査した場合には、ステップS207で決定された代表色をクラスタリングする（ステップS209）。ここでのクラスタリング手法は、よく知られた代表的な手法でよい。クラスタリングの結果から、ステップS210において最大クラスターの代表色（Bとする）を選択肢、Bを指定領域Aの背景色とする（ステップS211）。

【 0 0 6 4 】

図 5 は、縮小画像生成処理（ステップ S 1 0 3）を示すフローチャートである。縮小画像生成処理は、画像へのアクセスコストを押さえるために行う処理であり、原画像から縮小画像を生成して、ステップ S 1 0 4 以降の処理は縮小画像に対して行われる。

【 0 0 6 5 】

ステップ S 3 0 1 にて、オリジナル画像を入力し、ステップ S 3 0 2 にて、上記のようにして求めた背景色 B を入力する。次に、入力画像を  $m \times m$  の格子状に分割し（ステップ S 3 0 3）、左上から順に格子を選択する（ステップ S 3 0 4）。本実施例では、 $3 \times 3$  の格子サイズを使用する。なお、格子はブロックと称してもよい。

【 0 0 6 6 】

ステップ S 3 0 5 にて、格子  $i$  の代表色を  $C_i$  とする。 $C_i$  は、例えば格子内の各画素値の平均値とすることができる。次に、ステップ S 3 0 6 にて、格子  $i$  の色の分散値を  $V_i$  とする。次に、格子  $i$  の代表色  $C_i$  と背景色 B とを比較し、また、格子  $i$  の色の分散値  $V_i$  と所定の値とを比較する（ステップ S 3 0 7）。比較の結果、格子  $i$  の代表色  $C_i$  が背景色 B とほぼ同じであり、かつ、格子  $i$  の色の分散値  $V_i$  が十分小さければ（ステップ S 3 0 7 における YES）、その格子  $i$  の代表色を B と定義して、格子を 1 つの画素値 B に圧縮する（ステップ S 3 0 9）。ステップ S 3 0 7 において NO の場合、格子  $i$  の代表色をそのまま  $C_i$  と定義して、上記と同様に、格子を 1 つの画素値  $C_i$  に圧縮する（ステップ S 3 0 8）。すなわち、 $1/9$  に圧縮される。ステップ S 3 1 0 にて全ての格子を処理したかどうかを調べ、全ての格子を処理したならば処理を終了し、まだ全て処理していなければ格子の選択（ステップ S 3 0 4）の処理から再び処理を行う。

【 0 0 6 7 】

上記の画像圧縮処理について更に説明する。図 6 は、画像圧縮処理（ $3 \times 3$  ブロック OR 圧縮と称する）の概念を 2 値画像を例にとり説明するための図である。同図の（a）に示すように、 $3 \times 3$  ブロックの画素が、例えば、黒背景上の白文字画素ならば、そのブロックを白文字画素で代表する。また、（b）に示すよ



うに、 $3 \times 3$  ブロックのある画素が、例えば、白背景上の黒文字画素ならば、そのブロックを黒文字画素で代表する。この 2 値画像の方法では、黒背景に薄い文字が書かれたような文書画像やカラー画像の場合、圧縮した時に薄い色の文字画素が消滅する可能性があるが、上述した本発明の方法によれば、そのような問題点は解決される。

## 【 0 0 6 8 】

本発明では、予め対象ブロックを含む広い近傍内で背景色 B を求めておく。そして、格子内の代表色が背景色とほぼ同じで、格子内の色分散が小さければ該当の格子は背景色で代表し、そうでなければ、格子の代表色で代表する。このような方法によって、文書要素は  $C_i$  として残り、それ以外は背景色 B となるので、カラー画像を効率良く圧縮することが可能となる。すなわち、隣接した文字は上記の OR 圧縮でつぶれて固まり、1mm 以上の間隔はそのまま保持されるので、領域識別の観点からは、ボトムアップ的に画素を統合して文書要素（矩形）を生成する本発明にとって最適な圧縮画像を得ることが可能となる。また、画像を平滑化する効果もある。

## 【 0 0 6 9 】

さて、画像の領域識別の処理においては、処理速度および使用メモリを考慮すると画像の圧縮は欠かせないが、領域識別処理自体には OCR 処理ほどの高解像度・高品質な画像は必要ない。従って、上記のような格子ブロック処理による画像圧縮は、実装および速度の点から有利であり、処理速度が向上し、使用メモリ量の軽減を図ることができる。

## 【 0 0 7 0 】

ここで、図と文字とは少なくとも 1mm 以上離れていると考えられるので、本実施例では、ブロックサイズを  $3 \times 3$  としている。すなわち、3 ドットの圧縮であれば、図と文字とは融合されずに済む。200 dpi の画像ならば  $3 \times 3$  又は  $4 \times 4$  が適当なブロックサイズであると考えられる。

## 【 0 0 7 1 】

次に、ラン生成の処理（ステップ S104）について更に説明する。

## 【 0 0 7 2 】

一般に 2 値画像の場合は、同一ライン上の近傍の黒画素同士を統合して黒ランを生成する。一方、本発明では、背景色 B とほぼ同じ画素値を持つ画素を 0、それ以外の画素を 1 として、1 の画素値を持つ画素を統合してランを生成する。背景色と注目画素値が同一かどうかの判断は、図 7 に示すフローチャートに示すような手順で行われる。すなわち、背景色 B と注目画素値 P を入力し（ステップ S 5 0 1、S 5 0 2）、ステップ S 5 0 3 にて、RGB 各色値の差の絶対値がそれぞれ同時に十分に小さいかどうかを調べる。同時に十分に小さい場合は B と P は同一色であると判断し（ステップ S 5 0 4）、そうでない場合は、B と P は異なる色であると判断する（ステップ S 5 0 5）。RGB 系で表した背景色 B を  $(r, g, b)$ 、ある画素 P を  $(r_p, g_p, b_p)$  とする場合、RGB 各色値の差の絶対値がそれぞれ同時に十分小さいときとは、例えば、B と P が次の条件を満たすときである。

【0 0 7 3】

$$|r - r_p| < 100 \text{ かつ } |g - g_p| < 100 \text{ かつ } |b - b_p| < 100$$

上記の方法を画像の左上から順に 1 行ずつ適用することにより、カラー画像からランを生成する。ランは例えば左上と右下の座標値  $(x_s, y_s, x_e, y_e)$  を持つ矩形である。

【0 0 7 4】

ラン統合・矩形生成処理（ステップ S 1 0 5）においては、現在生成中のランと、前ラインとのランとの間に重なりがある場合は、前ラインのランの座標値を更新する形で現在生成中のランと前ラインのランを統合して矩形を生成する。2 つの矩形が重なりを持つ場合には、2 つの矩形を包含する最小の座標値を求めて、一方の座標値を更新する形で統合する。

【0 0 7 5】

ここで、矩形（又は矩形データ）とは、矩形の左上と右下の座標  $(X_s, Y_s, X_e, Y_e)$ 、属性（文字、罫線、図・写真等）、及び任意属性（文字色、背景色等）からなる組である。矩形の例を次に示す。

【0 0 7 6】

写真矩形  $R1 = \{(20, 40, 64, 72), \text{“写真”}\}$

文字矩形  $R_2 = \{(20, 40, 30, 50), \text{“文字”}, \text{黒}, \text{白}\}$  第3, 4要素は文字色（黒）と背景色（白）である。ランも矩形の一種であり、ランの場合は、 $y_s = y_e$  である。

## 【 0 0 7 7 】

矩形統合・矩形リスト生成処理（ステップ S 1 0 6）における領域識別結果は、次のような矩形データのリストで表現される。

## 【 0 0 7 8 】

矩形リスト =  $\{R_1, R_2, \dots, R_n\}$ ,  $n$ : 抽出された矩形の総数

次に、矩形分類処理（ステップ S 1 0 7）について説明する。矩形分類処理においては、ステップ S 1 0 6 にて生成した矩形を文字、罫線、および図・写真（その他）、周辺ノイズのそれぞれに分類する処理を行う。この中で、本発明で特に重要なのは、図・写真（その他）である。分類の際の特徴には、例えば、矩形のサイズ、高さ、位置、縦横比、等を用いる。以下、図 8 のフローチャートを用いて矩形分類処理について説明する。

## 【 0 0 7 9 】

まず、生成された矩形を、生成された順にリストや配列として入力する（ステップ S 4 0 1）。続いて、矩形リストの先頭から最後まで順に以下の処理を行う。

## 【 0 0 8 0 】

現在参照中の矩形を  $R$  とおく（ステップ S 4 0 2）。ステップ S 4 0 3 にて、矩形  $R$  が罫線特徴を有するかどうかを調べる。矩形  $R$  が横または縦方向に十分細長ければ、 $R$  を横または縦方向の罫線と分類する（ステップ S 4 0 6）。ステップ S 4 0 3 で NO であれば、ステップ S 4 0 4 において、矩形  $R$  の高さが十分小さいかどうかを調べ、十分に小さければ矩形  $R$  は文字矩形であると分類する（ステップ S 4 0 7）。ステップ S 4 0 4 で NO の場合であって、矩形  $R$  の面積が十分に小さく場合も、矩形  $R$  は文字矩形であると分類する（ステップ S 4 0 5、S 4 0 7）。

## 【 0 0 8 1 】

ここで、S 4 0 3 における罫線特徴を持つかどうかは、具体的には次の条件を

満たすかどうかで判断する。すなわち、(1) 矩形の縦横比が 20 倍以上、(2) 高さが 3 ドット以下、(3) 矩形内部において長い水平(垂直)ランのみで構成されている、場合に罫線特徴を持つとする。

【0082】

また、ステップ S 4 0 4 において、R の高さが十分小さいとは、具体的には紙の上で 10 mm 以下とする。すなわち、200 dpi 画像の場合には高さ 80 ドット以下の矩形がこの条件に相当する。

【0083】

更に、ステップ S 4 0 5 において、R の面積が十分小さいとは、具体的には 1600 (dot×dot) 以下とする。

【0084】

ステップ S 4 0 3 ~ S 4 0 5 のいずれの条件にも当てはまらなかった矩形は、ステップ S 4 0 8 において、一旦図・写真(その他)矩形の候補として分類し、以降の処理で、図・写真(その他)の候補に分類された矩形をさらに詳しく分類する。

【0085】

すなわち、図・写真(その他)矩形を1つの文書画像とみなして、本発明の領域識別処理を再帰的に施す。すなわち、ステップ S 4 0 9 で矩形 R の背景色を特定し、ステップ S 4 1 0 ~ S 4 1 2 で矩形 R 内の矩形を生成する。ただし、矩形 R はすでに原画像から圧縮されているので、R を再び圧縮する必要はない。従って、図 2 のステップ S 1 0 3 に相当する処理は行わない。

【0086】

ステップ S 4 1 3 において、R に対して再帰的に図 8 に示したフローチャートの矩形分類処理を施す。ここで、再帰的に矩形生成・分類処理を行う理由は次の通りである。

【0087】

2 値画像と異なり、カラー画像では、領域毎に異なる背景色を持つ場合がある。そのような場合、文書全体の背景色 B と異なる背景色を持つある部分領域について、上記の矩形生成の結果、その部分領域全体が誤って1つの矩形として統合

されてしまうことがある。従って、再帰的に矩形生成・分類処理を行う必要がある。

#### 【 0 0 8 8 】

例えば、図 9 におけるオリジナル画像の背景色 1 と部分領域 3 内の背景色 5 が異なる場合、領域識別結果として、領域 3 は 1 つの矩形として識別される。この領域 3 に再び矩形生成・分類処理を施すこととなる。

#### 【 0 0 8 9 】

ステップ S 4 1 4 にて、R 内の文字矩形を抽出するために、R 内の同一ライン上で連続した文字矩形が存在するかどうかを調べる。存在しない場合、R を正式に図・写真などのその他矩形に分類して（ステップ S 4 1 6）、矩形リストの次要要素を取り出して（ステップ S 4 2 0）、次の矩形の処理に移る（ステップ S 4 2 1）。ここで、矩形リスト中の全ての矩形を分類し終われば処理を終了する（ステップ S 4 2 1 で Y E S の場合）。

#### 【 0 0 9 0 】

ステップ S 4 1 4 において、R 内の同一ライン上で連続した文字矩形が存在する場合、ステップ S 4 1 5 にて、これら文字矩形を包含するような大きな文字矩形 T を生成する。R 内の文字部分がいくつか離れた箇所に複数存在する場合には、上記のようにして統合した矩形 T も R 内に複数生成されると考えられる。次に、ステップ S 4 1 7 において、矩形 T を文字矩形として矩形リスト L に追加登録する。

#### 【 0 0 9 1 】

ここで、ステップ S 4 1 4 において、R 内に同一ライン上で連続した文字矩形が存在するかどうかは、具体的には、先頭文字矩形の高さの中心を基準ラインとしたとき、基準ラインの  $\pm 5 \text{ dot}$  以内に連続した文字矩形が 3 つ以上連続しているかどうかで判断する。

#### 【 0 0 9 2 】

続いて、ステップ S 4 1 8 にて、T の面積が R の面積に対して十分大きいかどうかを調べる。十分大きい場合には、ステップ S 4 1 9 にて、R を矩形リストから削除する。ステップ S 4 1 8 にて N O の場合は、R を正式に図・写真（その他

）矩形に分類して（ステップ S 4 1 6）、次の矩形の処理に移る（ステップ S 4 2 1）。以上の処理をすべての矩形について繰り返す。

【 0 0 9 3 】

ここで、ステップ S 4 1 8において、Tの面積がRの面積に対して十分大きいかどうかは、具体的にはTの面積がRの面積の80%以上であるかどうかで判断する。

【 0 0 9 4 】

なお、上記の矩形分類処理において、ステップ S 4 0 1 から S 4 0 8 までの処理は2値画像に対する領域識別処理と同様な処理である。カラー画像特有の処理は、ステップ S 4 0 8 でその他と分類された矩形に対して、再帰的にもう一度図 8 の処理を施すことである。

【 0 0 9 5 】

次に、写真矩形を塗り潰した画像を生成する処理（ステップ S 1 0 9）について、図 1 0 の例を用いて説明する。

【 0 0 9 6 】

図 1 0 の左上にオリジナル画像 1 0、右上に領域分割画像 1 2 を示す。領域分割画像 1 2 には、本発明のカラー領域識別の結果、写真矩形と判断された矩形を特定色で塗り潰す処理（ステップ S 1 0 9 の処理）を施してある。

【 0 0 9 7 】

入力カラー画像を2値化してOCR処理を施す場合、写真矩形を塗り潰さずにそのままの状態画像全体を2値化して、生成された2値画像から文字行矩形を生成した場合、図 1 0 の左下のような画像 1 1 となる場合がある。すなわち、写真矩形の薄い部分の近傍の文字矩形が、誤って写真部分と統合されてしまう。一方、ステップ S 1 0 9 の処理を施した領域分割画像 1 2 から文字行を生成した結果の画像 1 3 では、正しく文字矩形が生成される。

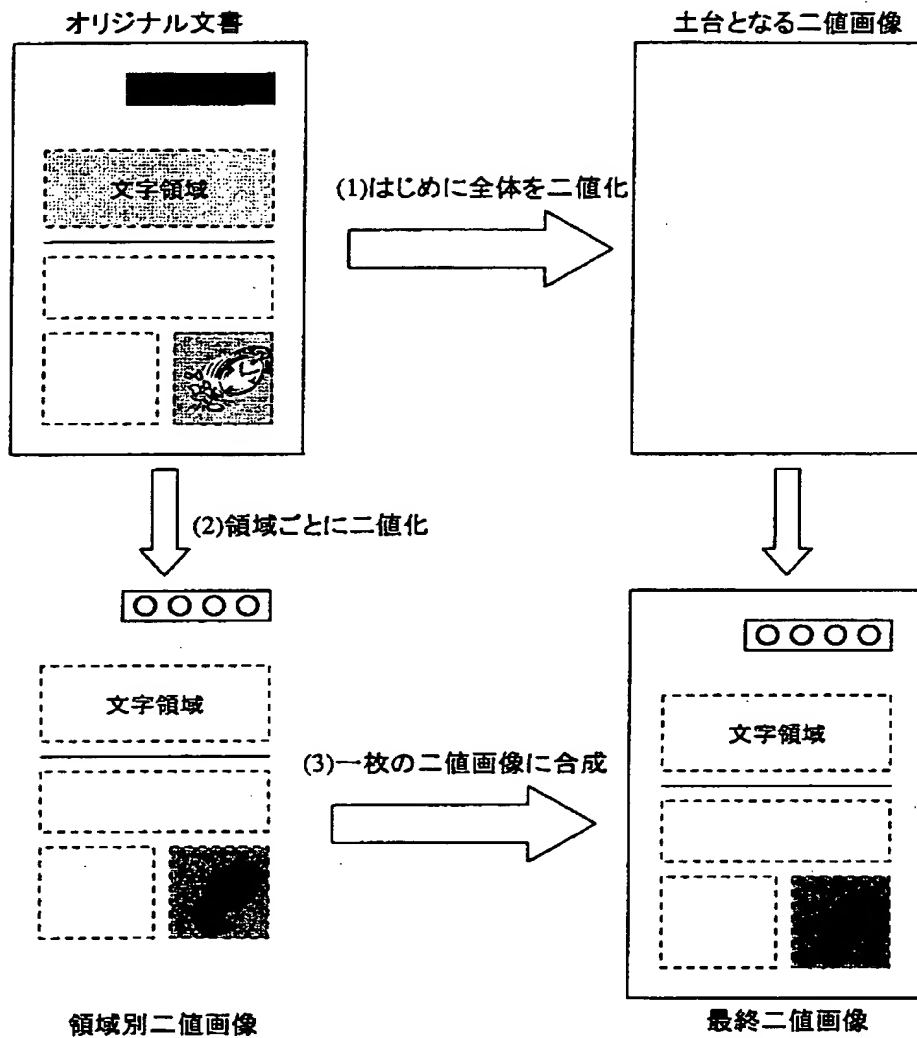
【 0 0 9 8 】

本発明の処理によって得られた領域識別結果はOCR以外の画像認識処理に適用することも可能である。

【 0 0 9 9 】

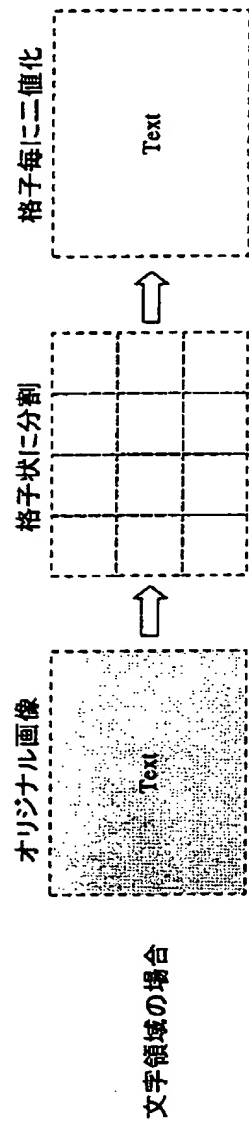
【図 1 5】

本発明の第 2 の実施例における二値化方法を説明するための図



【図 1 6】

画像分割型二値化方法を説明するための図





【図 1 7】

領域の幅に対する格子幅を示す表

領域の幅	格子の幅
512	32
1024	64
2048以上	128

【図 1 8】

領域識別の結果例

オリジナル画像



カラー領域識別結果

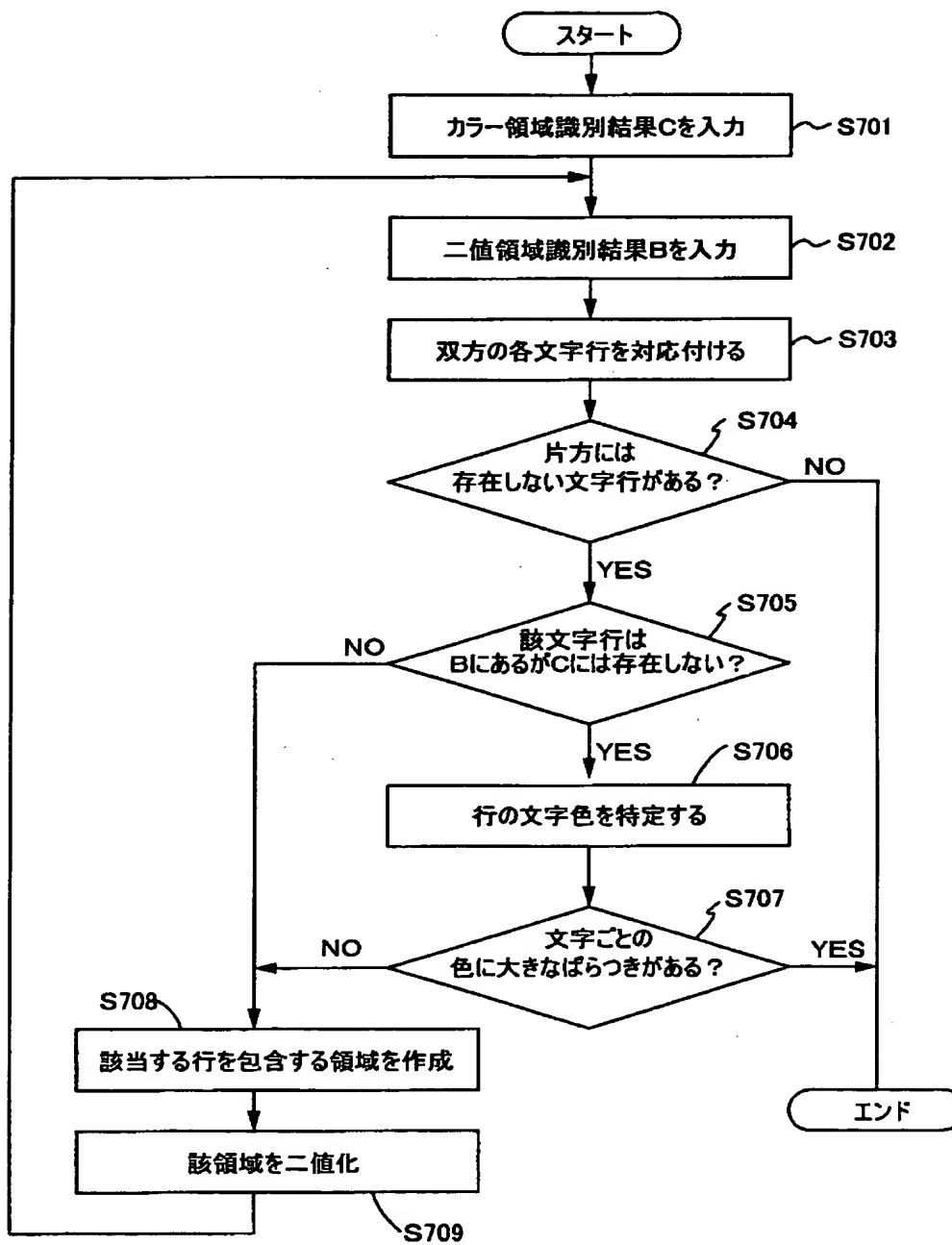


二値領域識別結果



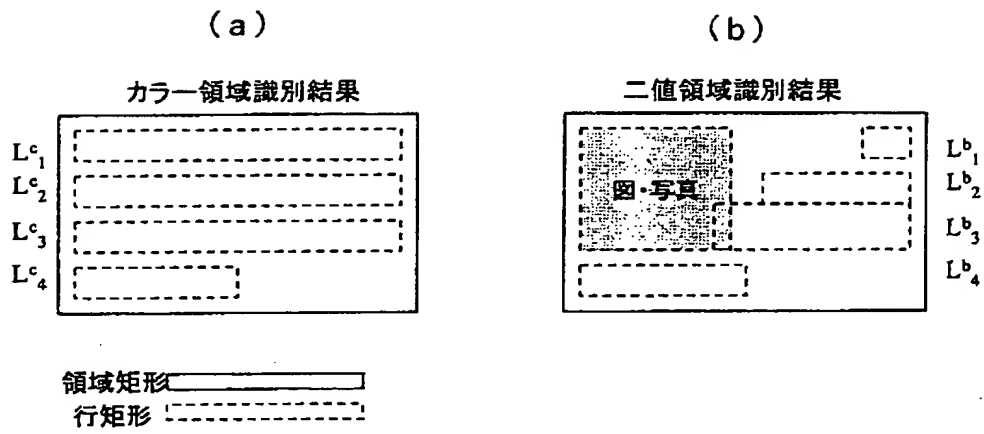
【図 1 9】

図18に示す場合において領域識別結果の修正を行う処理を示すフローチャート



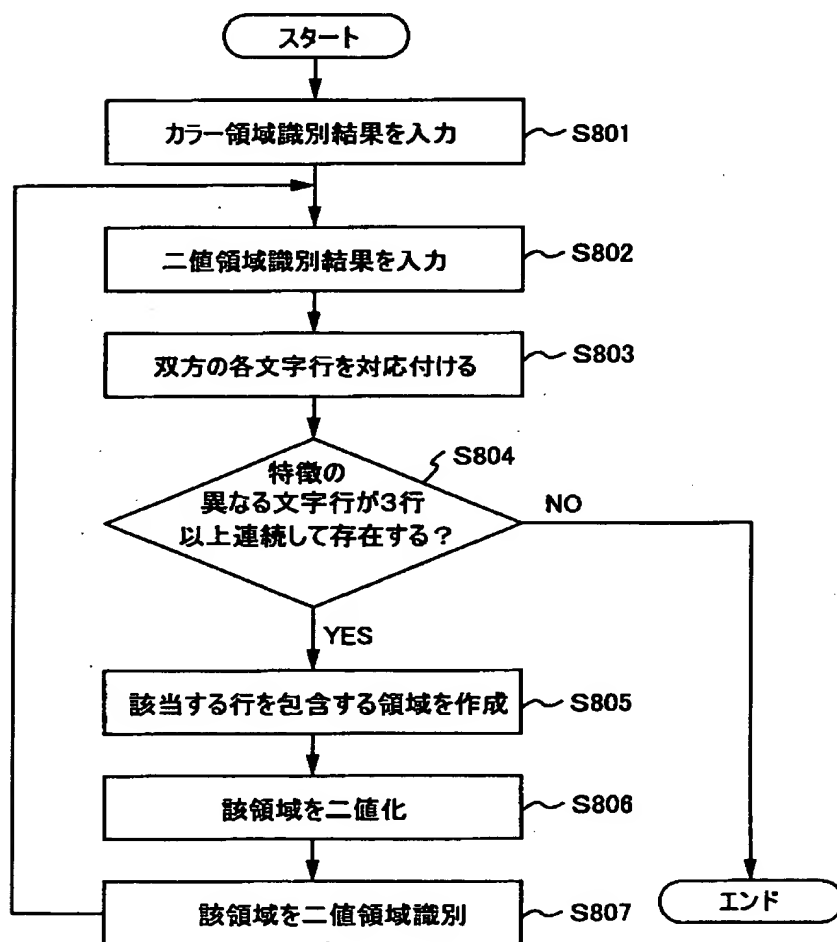
【図 2 0】

領域識別の結果例



【図 2 1】

図20に示す場合において領域識別結果の修正を行う処理を示すフローチャート



【書類名】 要約書

【要約】

【課題】 カラー文書画像及び白黒・グレー画像の領域を精度良く、効率的に識別し、また、カラー文書特有の問題を有するカラー文書に対しても精度よくOCRすることができる文書画像認識方法を提供する。

【解決手段】 文書画像を認識する文書画像認識方法が提供される。その方法では、文書画像をデジタル画像として入力し、該文書画像の背景色を特定し、必要に応じて画像を縮小し、該背景色を用いて該文書画像から背景領域以外の画素を抽出し、該画素を統合して連結成分を生成し、該連結成分を少なくとも形状特徴を用いて所定の領域に分類して、該文書画像の領域識別の結果を得る。また、二値画像の領域識別を行い、その結果とカラー領域識別結果とを照合し、必要に応じてフィードバック処理を行ってOCRに好適な二値画像と領域識別結果を得るようにする。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号 [000006747]

1. 変更年月日 1990年 8月24日

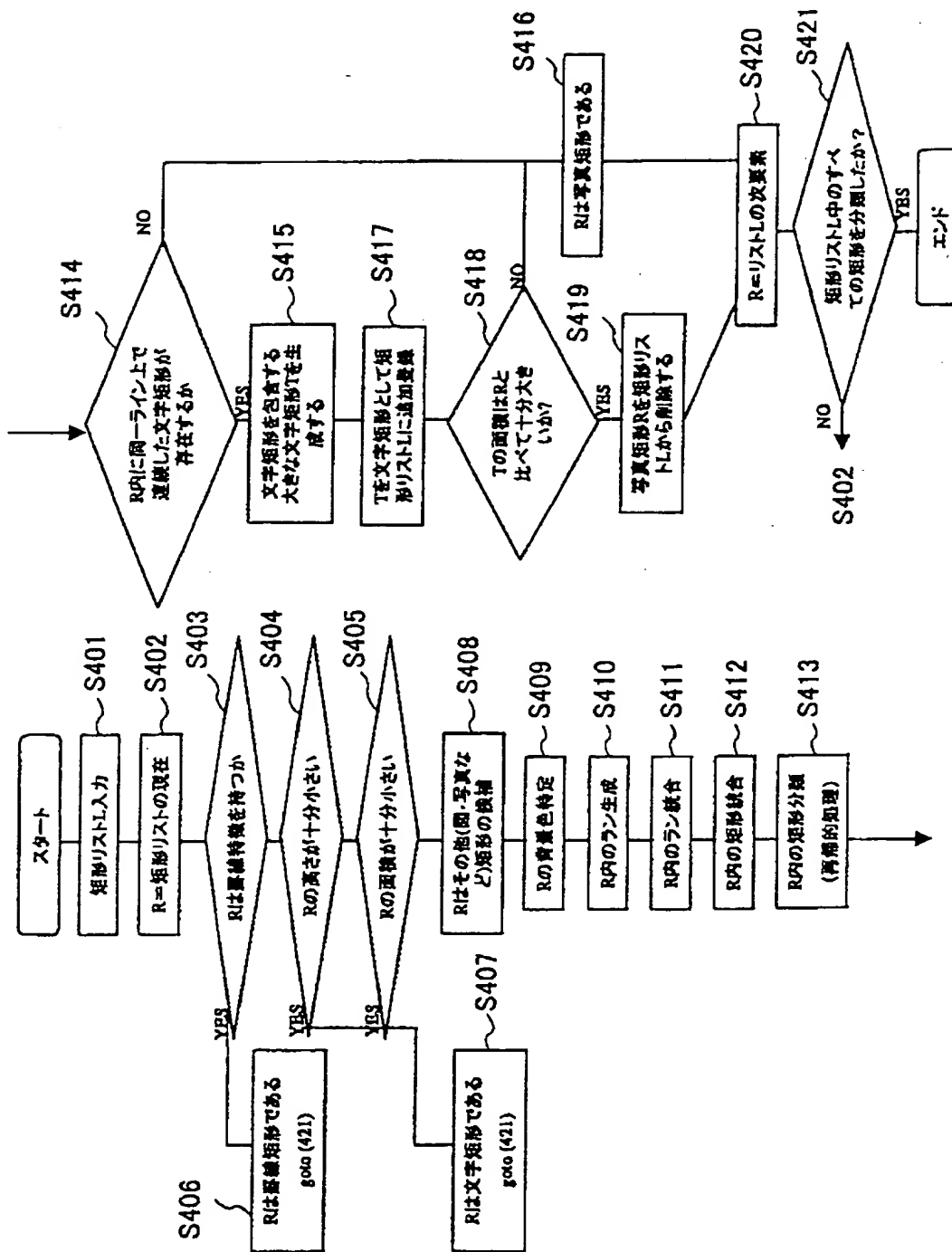
[変更理由] 新規登録

住 所 東京都大田区中馬込1丁目3番6号

氏 名 株式会社リコー

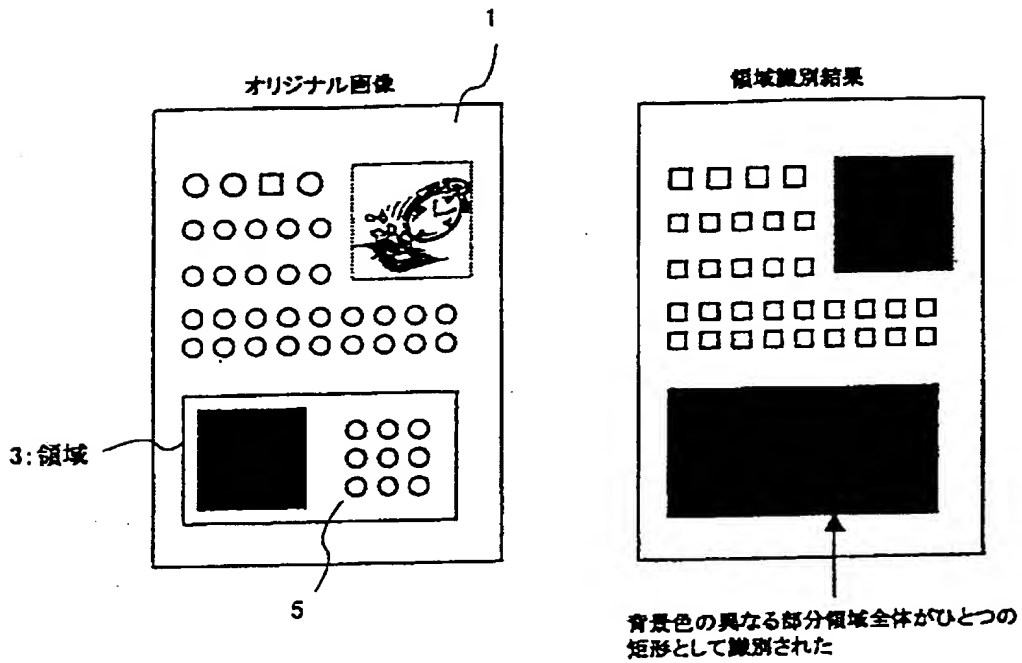
【図 8】

矩形分類処理を示すフローチャート



【図 9】

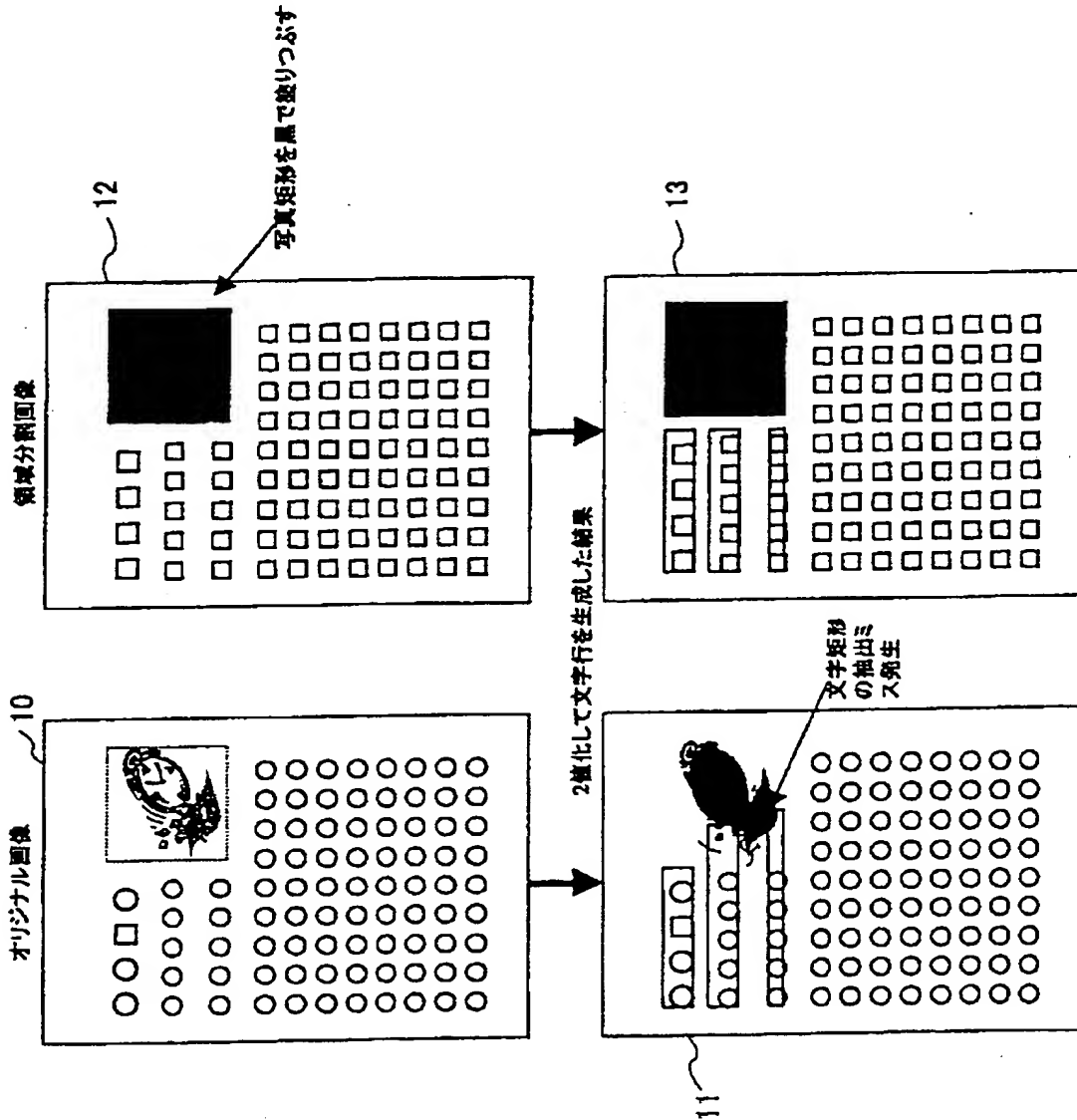
部分領域全体が矩形として識別された例を示す図





【図 1 0】

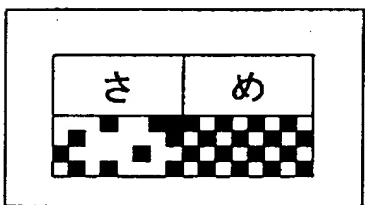
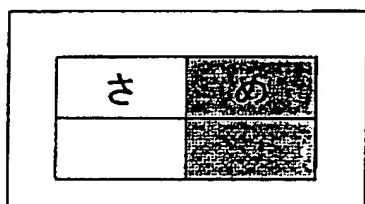
写真矩形を塗り潰した画像を生成する処理を説明するための図



【図 1 1】

各セルが色分けされた表の例

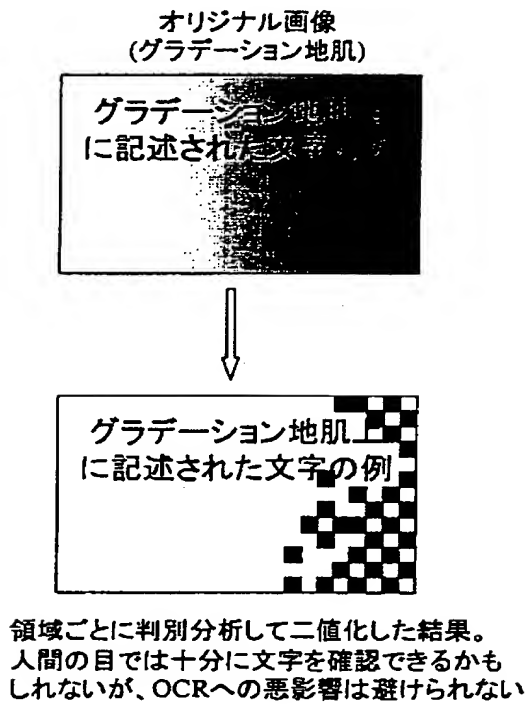
オリジナル画像  
(各セルが色分けされた表)



領域ごとに判別分析して二値化

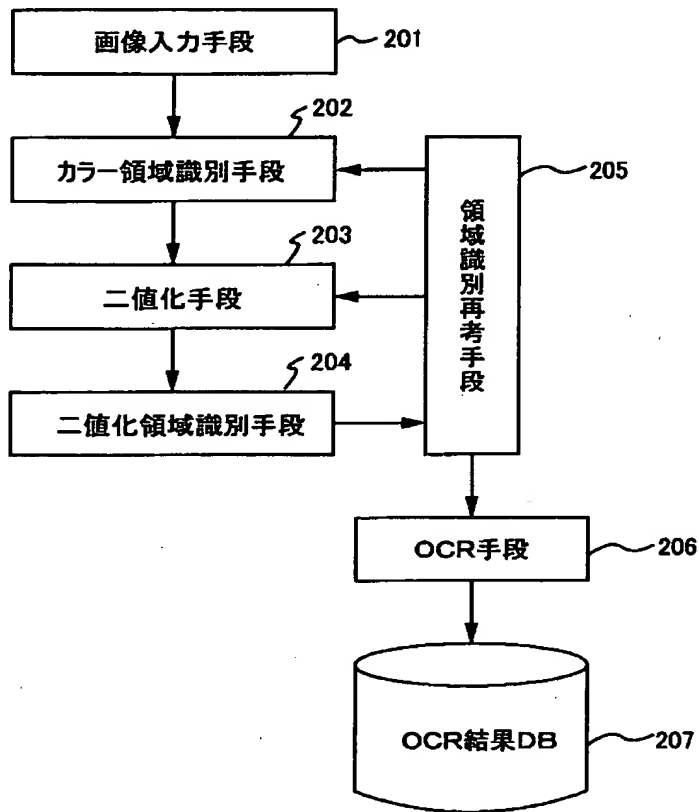
【図 1 2】

グラデーション地肌に記述された文字の例を示す図



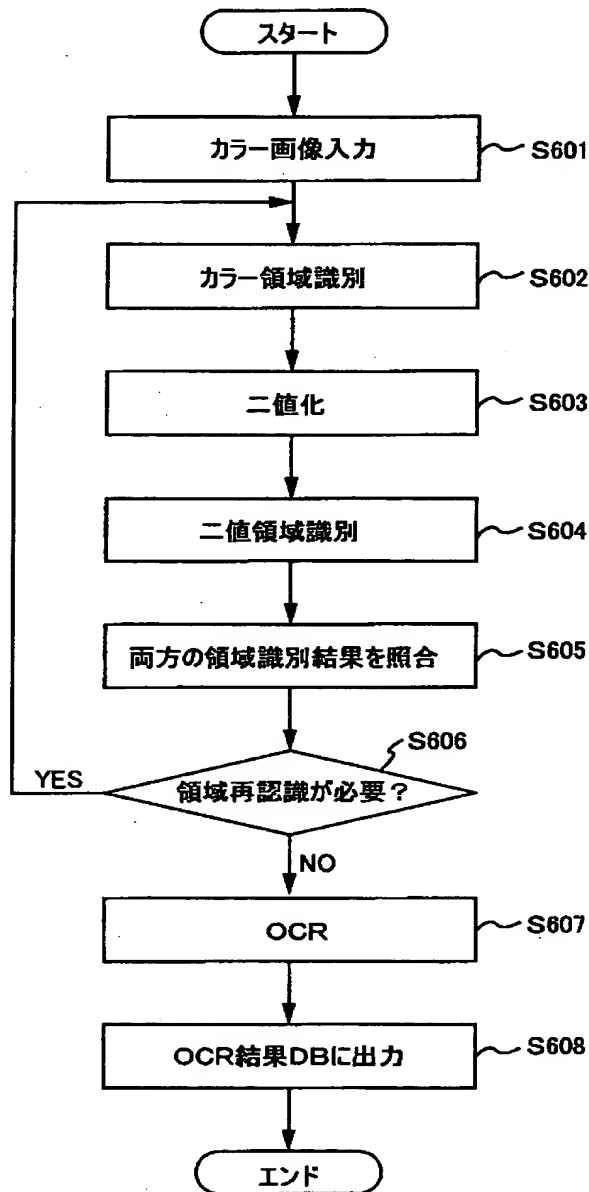
【図 1 3】

本発明の第2の実施例における文書画像認識装置の構成図



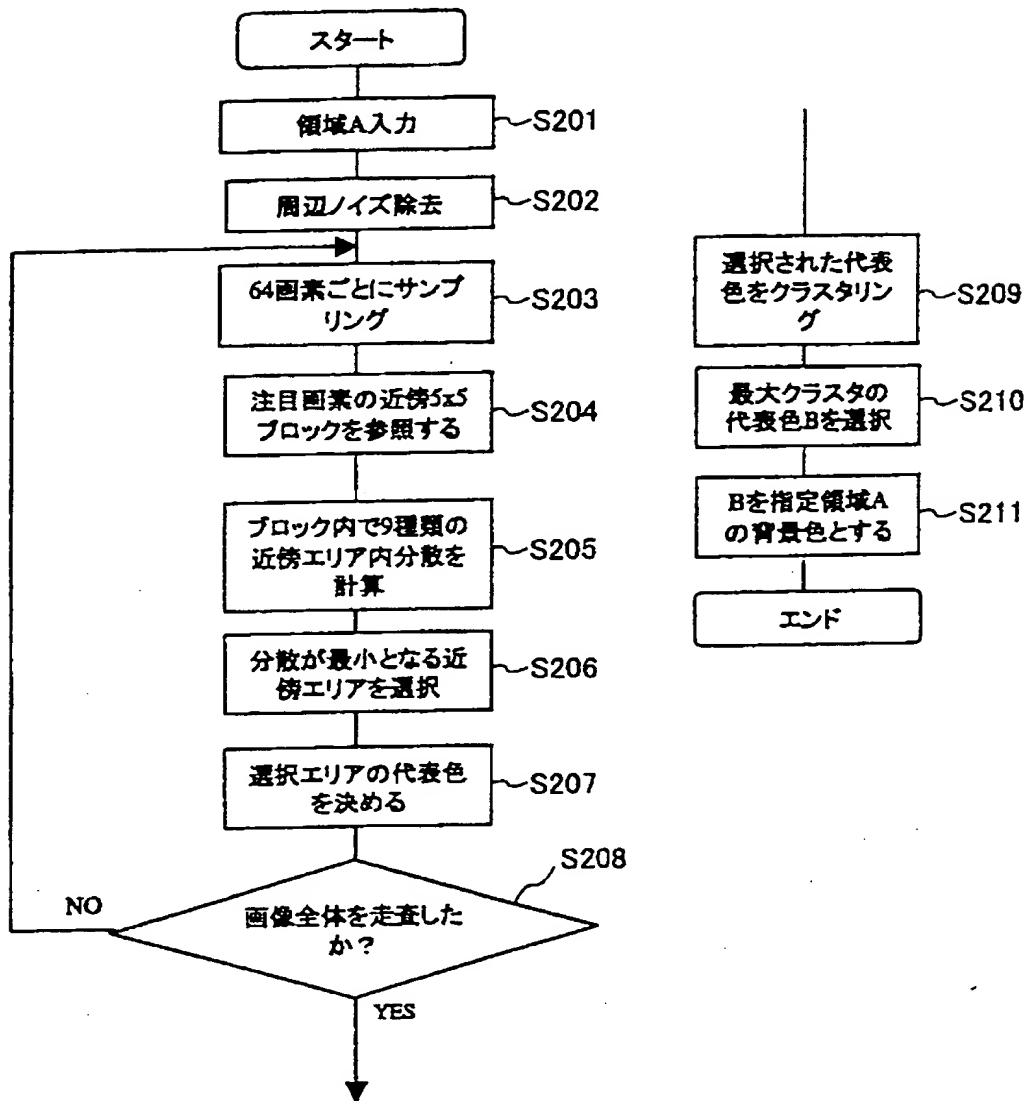
【図 1 4】

本発明の第2の実施例における文書画像認識装置の動作を示すフローチャート



【図 3】

背景色特定処理を示すフローチャート



【図 4】

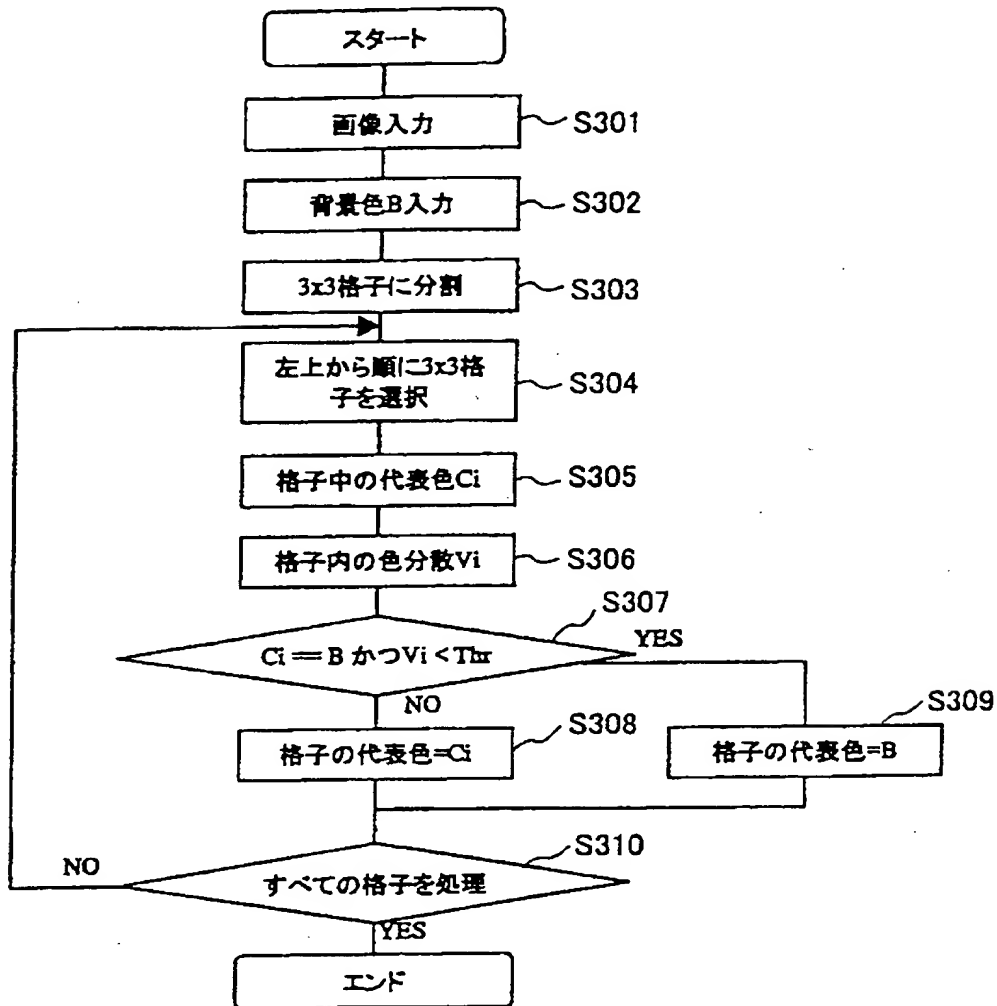
近傍 5 × 5 ブロック内を参照した平滑化における  
9 種類のエリアの例を示す図

area0	area1	area2	area3	area4	area5	area6	area7	area8
00...	...11	.....	.....	.444.	.....	.....	.....	.....
000..	..111	.....	.....	.444.	55...	...66	.....	.888.
.0P..	..P1.	.2P..	..P3.	..P..	55P..	..P66	..P..	.8P8.
.....	.....	222..	..333	.....	55...	...66	.777.	.888.
.....	.....	22...	...33	.....	.....	.....	.777.	.....

各エリア (area0, area1, ..., area8) は中心画素 p と n で表わされる画素で構成される  
 p : 5 × 5 の中心点  
 . : エリア外の点  
 n : エリア n (=0...8) の点

【図 5】

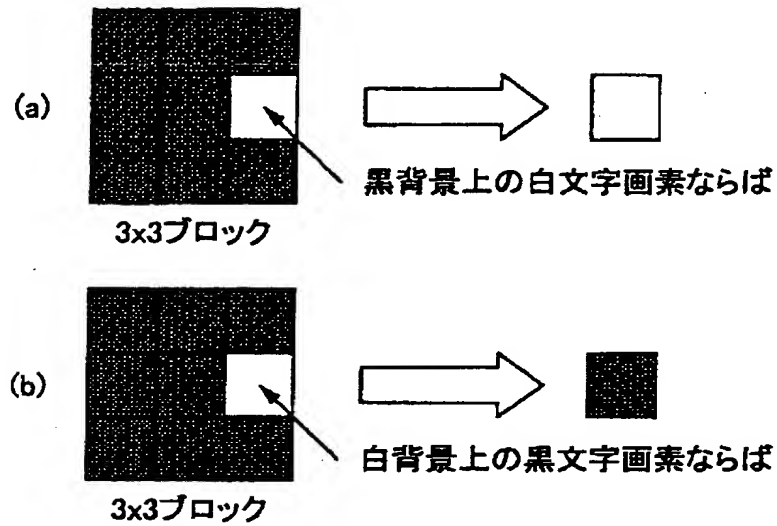
縮小画像生成処理を示すフローチャート





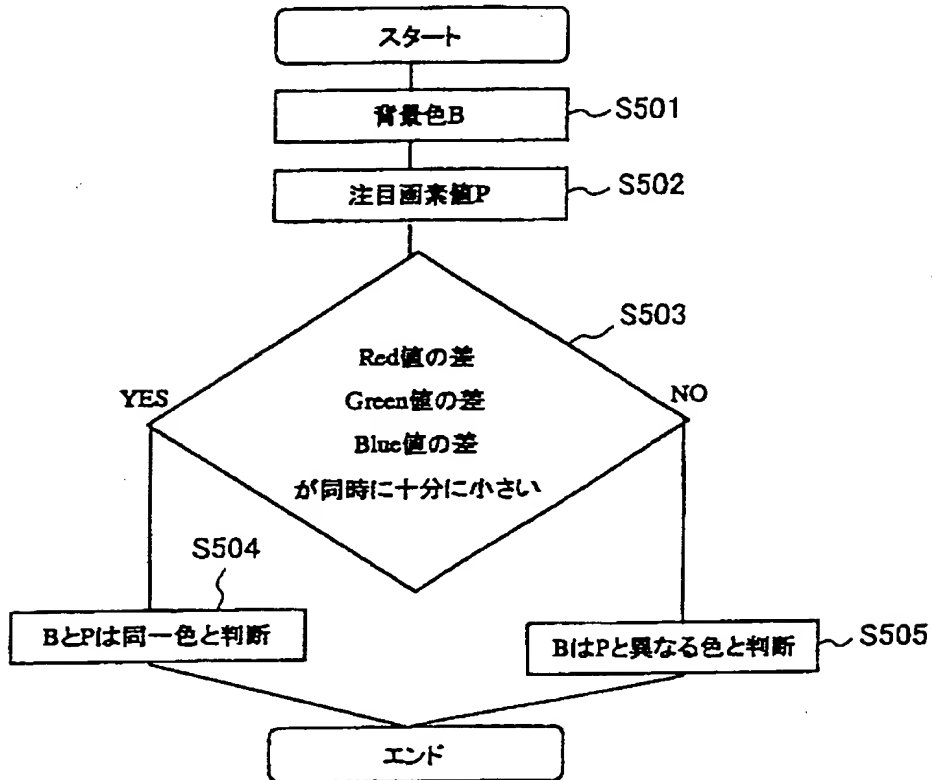
【図 6】

縮小画像生成処理の概念を 2 値画像を例にとり説明するための図



【図 7】

背景色と注目画素値が同一かどうかの判断を行う  
処理を示すフローチャート



更に、2値化の影響を受けやすい図・写真（その他）矩形を特定色で塗り潰すこととしたので、図・写真（その他）領域の周辺にある文字部を精度よく抽出することができ、OCR精度も向上する。

## 【0154】

また、第2の実施例で説明したように、OCR前処理（領域識別、二値化）へのフィードバック機構を有するので、従来のカラー文書認識技術では難しかったカラー文書特有の問題を有するカラー文書に対しても精度よくOCRすることができる。また、各セルが色分けされた表や、グラデーションが大きく変化する地肌に書かれた文字を有するカラー文書、あるいはこれらが混在する場合でも精度よくOCRすることができる。

## 【0155】

更に、領域単位ではなく文字行単位で領域識別結果を比較することにより、従来避けられなかった行から領域への統合誤りを回避することができる。その際、文字色を特定して文字色のばらつきを調べることで、比較対象範囲に文字が存在するか否かを精度よく判定でき、その結果文字が存在する範囲だけをフィードバックすることで周囲への悪影響がなく文字抽出精度が向上する。また、カラー情報を積極的に用いるカラー領域識別での文字抽出結果を優先的に採用することで、二値化が原因で従来OCRできなかった文字が抽出可能となる。更に、対象範囲の複数行についてそれぞれレイアウト特徴を比較することで、どちらかの領域識別結果での誤りを正確に検出することができ、該当範囲だけをフィードバックすることで周囲への悪影響がなく文字抽出精度が向上する。

## 【図面の簡単な説明】

## 【図1】

本発明の第1の実施例における文書画像認識装置の構成図である。

## 【図2】

本発明の第1の実施例における処理の全体を示すフローチャートである。

## 【図3】

背景色特定処理を示すフローチャートである。

## 【図4】

近傍 5 × 5 ブロック内を参照した平滑化における 9 種類のエリアの例を示す図である。

【図 5】

縮小画像生成処理を示すフローチャートである。

【図 6】

縮小画像生成処理の概念を 2 値画像を例にとり説明するための図である。

【図 7】

背景色と注目画素値が同一かどうかの判断を行う処理を示すフローチャートである。

【図 8】

矩形分類処理を示すフローチャートである。

【図 9】

部分領域全体が矩形として識別された例を示す図である。

【図 1 0】

写真矩形を塗り潰した画像を生成する処理を説明するための図である。

【図 1 1】

各セルが色分けされた表の例である。

【図 1 2】

グラデーション地肌に記述された文字の例を示す図である。

【図 1 3】

本発明の第 2 の実施例における文書画像認識装置の構成図である。

【図 1 4】

本発明の第 2 の実施例における文書画像認識装置の動作を示すフローチャートである。

【図 1 5】

本発明の第 2 の実施例における二値化方法を説明するための図である。

【図 1 6】

画像分割型二値化方法を説明するための図である。

【図 1 7】

領域の幅に対する格子幅を示す表である。

【図 1 8】

領域識別の結果例である。

【図 1 9】

図 1 8 に示す場合において領域識別結果の修正を行う処理を示すフローチャートである。

【図 2 0】

領域識別の結果例である。

【図 2 1】

図 2 0 に示す場合において領域識別結果の修正を行う処理を示すフローチャートである。

【符号の説明】

- 1 オリジナル画像の背景色
- 3 部分領域
- 5 部分領域 3 の背景色
- 1 0 オリジナル画像
- 1 1 オリジナル画像 1 0 の文字行生成結果
- 1 2 領域分割画像
- 1 3 領域分割画像 1 2 の文字行生成結果
- 1 0 0 C P U
- 1 0 1 メモリ
- 1 0 2 スキャナ
- 1 0 3 通信装置
- 1 0 4 表示装置
- 1 0 5 ハードディスク
- 1 0 6 キーボード
- 1 0 7 C D - R O M ドライブ
- 2 0 1 画像入力手段
- 2 0 2 領域識別手段

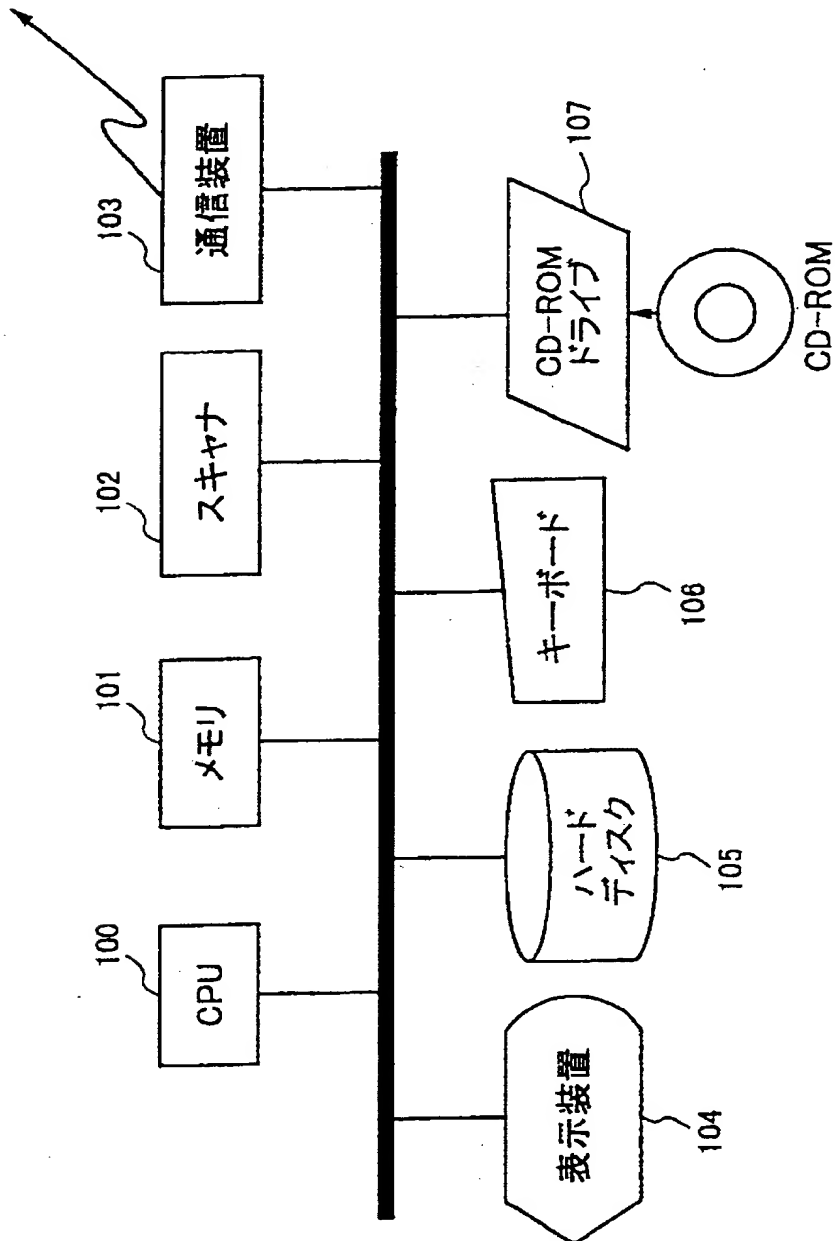
- 2 0 3 二値化手段
- 2 0 4 領域識別手段
- 2 0 5 領域識別再考手段
- 2 0 6 OCR手段
- 2 0 7 データベース

【書類名】

図面

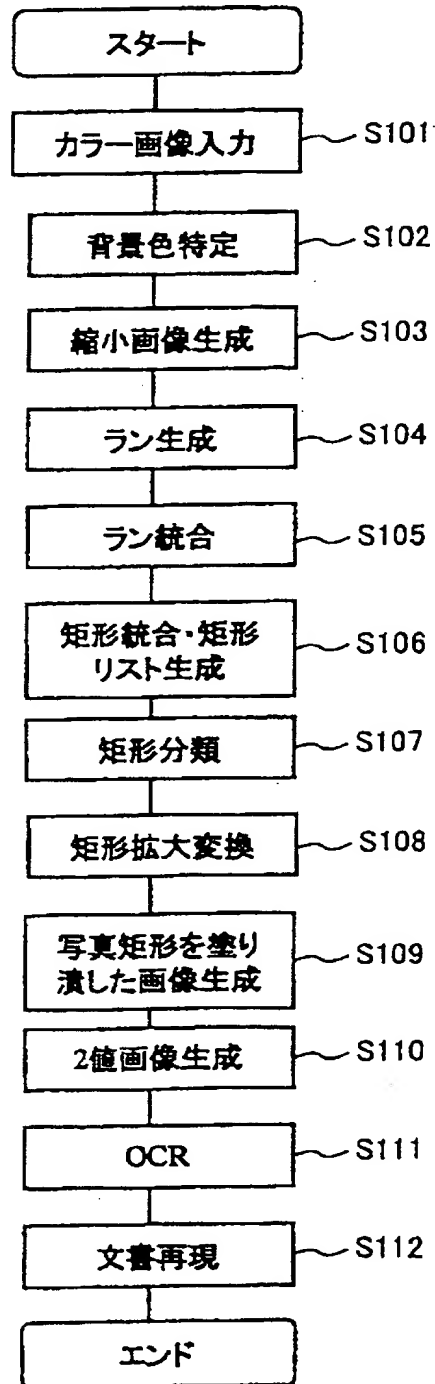
【図 1】

本発明の第 1 の実施例における文書画像認識装置の構成図



【図 2】

本発明の第 1 の実施例における処理の全体を示すフローチャート





## (第 2 の実施例)

次に、本発明の第 2 の実施例について説明する。

## 【 0 1 0 0 】

前述したように、一般にカラー化された文書では、単に使用される色数が増えるだけではなく文書のレイアウト構造も複雑になる傾向がある。このため、単純に二値化して領域識別しても、最終的な OCR 精度を向上させるのは難しい。

## 【 0 1 0 1 】

例えば、図 1 1 に示すような各セルが色分けされた表の場合、従来の二値化技術を用いると表として適切に二値化されない場合がある。また、図 1 2 に示すようなグラデーションが大きく変化する地肌に書かれた文字の場合にも、OCR 処理のために適切な画像が得られない場合がある。

## 【 0 1 0 2 】

ここで、個々の文字領域の背景は一定の色あるいは緩やかなグラデーションであると仮定すれば、第 1 の実施例で説明した方法を用いることによって対応可能である。また、文字は必ず一定の色で記述されていると仮定すれば、画像分割型の二値化方式（特願平 1 1 - 1 1 3 7 6 1 号参照）でも OCR 可能なレベルの状態にはなる。しかし、一般にカラー化された文書は複雑であり、図 1 1 や図 1 2 に示すようなカラー特有の問題が混在している場合が多い。従って、一般にカラー化された文書に対しては、従来方法（二値化＋領域識別）を組み合わせるだけでは OCR に好適な画像を生成することは困難であるといえる。

## 【 0 1 0 3 】

第 2 の実施例では、第 1 の実施例で説明したカラー領域識別処理に加え、OCR に好適な画像を生成するために前処理へのフィードバック処理を追加して、個々の技術では対応しきれなかった複雑なカラー文書の認識を可能とした文書画像認識技術について説明する。第 2 の実施例では、カラー領域識別は文字よりも写真や図の抽出に強く、それに対して 2 値の領域識別はテキストの抽出に強いという性質を利用している。

## 【 0 1 0 4 】

図 1 3 に、本発明の第 2 の実施例における文書画像認識装置の構成を示す。な

お、第2の実施例における処理を実行するプログラムを用いて図1に示すコンピュータを第2の実施例における文書画像認識装置として使用することもできる。また、そのプログラムをCD-ROM等の記録媒体に記録して、コンピュータにインストールすることにより本発明を実施することもできる。

#### 【0105】

図13に示す文書画像認識装置は、画像入力手段201、領域識別手段202、二値化手段203、領域識別手段204、領域識別再考手段205、OCR手段206、データベース207を有する。

#### 【0106】

画像入力手段201により、紙文書であるカラー文書画像を入力する。例えば、スキャナー、あるいはネットワークを介してカラー文書画像をカラーデジタル画像として入力する。

#### 【0107】

領域識別手段202は、カラー文書画像を写真／テキスト／罫線などの小領域に分割する。ここでは第1の実施例におけるカラー領域識別処理と同様の処理を行う。

#### 【0108】

二値化手段203は、カラー文書画像を後段のOCRにとって最適な二値画像に変換する。後述するようにここでの二値化処理は、テキストと背景を精度よく分離し、罫線や写真などは擦れないように工夫され必要ならば文字のエッジを強調することも可能である。また、二値化手段203は性質の異なる複数の二値化方式を有しており、カラー領域識別で分類された各領域（写真／テキスト／罫線）をそれぞれ最適な方式で二値化する。

#### 【0109】

領域識別手段204は、二値画像に対して領域識別を行う。この領域識別については従来の領域識別方法を用いて行うことが可能である。

#### 【0110】

領域識別再考手段205は、カラー領域識別結果と二値領域識別結果との間に整合が取れているかどうかを確認して、もし両者の一部に違いがあれば該当する

領域のみを再度カラー領域識別、二値化、二値領域識別するようなフィードバック処理の判定を行う。このフィードバックは、両者の領域識別結果が一定の条件を満たすまで繰り返し実行される。OCR手段206は、二値画像を入力してOCR処理を行う。データベース207は、OCR結果を保持する。

## 【0111】

次に、第2の実施例における文書画像認識装置の動作を図14のフローチャートを用いて説明する。

## 【0112】

まず、ステップS601においてカラー画像を入力し、ステップS602にてカラー領域識別を行い、ステップS603にて識別された領域毎に異なる方式で二値化し、1枚のOCR用二値画像を生成する。

## 【0113】

そして、ステップS604において二値画像に対して従来方法で領域識別を行い、ステップS605にてカラー領域識別結果と二値領域識別結果とを比較し、ステップS606においてフィードバック処理が必要かどうかを判断する。

## 【0114】

判断の結果がYesならば、該当する領域だけ再度ステップS602からやり直し、判断の結果がNoならば、OCR処理を施して結果をDBに出力する（ステップS607、ステップS608）。

## 【0115】

上記のステップS606において、照合による矛盾が所定の値以下となった場合、もしくは所定回数だけS602～S605の処理を実行した場合にステップS607の処理に進むようにすることができる。また、S602～S605の処理結果がその前の処理結果と一致した場合にステップS607の処理に進むようにしてもよい。

## 【0116】

上記の処理のうちステップS603の領域ごとに二値化する処理と、ステップS606、606のフィードバック処理以外は第1の実施例で説明した処理と同様の処理を行う。

## 【 0 1 1 7 】

上記のように、カラー領域識別（ステップ S 6 0 2）において、入力されたカラー文書画像を二値化する前に色やレイアウト情報を用いて領域識別する。第 2 の実施例でも第 1 の実施例で説明した方法を用いることができる。前述した通り、この方法ではカラー文書画像を領域識別すると同時に、識別された文字矩形の文字色を検出することが可能である。第 2 の実施例では OCR 処理の前に文字色を利用するので、第 1 の実施例で説明したカラー領域識別方法は第 2 の実施例におけるカラー領域識別に適した方法である。

## 【 0 1 1 8 】

次に、領域ごとに二値化する処理（ステップ S 6 0 3）について詳細に説明する。

## 【 0 1 1 9 】

OCR にとって最適な二値画像は、文字や罫線、図・写真などの種類によって異なり、例えば、文字と背景とは明確に区別される必要がある。また文字の中に一部白抜けがあると、OCR への悪影響は避けられない。また、わずかな罫線のかすれも許されない。一方、図・写真領域は過剰分割を防ぐために全体としてぼんやりと暗めに二値化される方が望ましい。

## 【 0 1 2 0 】

以上の理由から、本実施例では図 1 5 に示すように文書の領域ごとに二値化して、OCR のための最適な二値画像を生成する。

## 【 0 1 2 1 】

図 1 5 の左上の画像はカラー領域識別された結果を矩形表示したものである。まず（1）に示すように、入力画像全面から最終二値化画像の土台となる二値画像を生成する。この二値化を行う方法は、判別分析法でも画像分割型二値化方式でも構わない。また、その他の方法で二値化してもよい。

## 【 0 1 2 2 】

次に、（2）に示すように、識別された領域ごとに各領域に適した方式で二値化が行われる。ここで、本実施例では、文字領域には画像分割型を、罫線、図・写真領域には判別分析法を適用する。

## 【0 1 2 3】

画像分割型二値化方法（特願平 1 1 - 1 1 3 7 6 1 号）は、図 1 6 に示すように、オリジナル画像を格子状に分割し、格子毎に二値化する方法である。この方法は、図 1 2 に示したようなかなりきつめ条件（グラデーション、周辺ノイズなど）でも、人間が認識可能なレベルで文字と背景を分離することが可能であるので、文字領域の二値化に適している。本実施例では、二値化対象領域の幅に応じて、格子幅を図 1 7 に示すように変化させることとしている。

## 【0 1 2 4】

一方、判別分析法で罫線、図・写真領域を二値化することで、特に二値化の閾値を濃い目に設定することで罫線がかすれず図や写真も一塊になり、領域識別にとって有利な二値画像となる。最後に、領域ごとに生成された二値画像を土台となる二値画像上に貼り付けることにより二値画像を生成する。

## 【0 1 2 5】

次に、図 1 4 のステップ S 6 0 5、S 6 0 6 で実行される領域識別結果の比較について詳細に説明する。

## 【0 1 2 6】

さて、例えば図 1 1 に示すように各行が色分けされた表を 2 値化処理した場合、表としてうまく 2 値化されない場合が多い。文書のカラー化に伴い、レイアウトもより一層複雑化する傾向があるので、2 値化処理のみではうまく 2 値化されない場合が一層多くなると考えられる。

## 【0 1 2 7】

そこで、本実施例においては、カラー領域識別結果と二値領域識別結果を比較して、必要ならばフィードバック処理を行って OCR のための最適な二値画像と正確な領域識別結果を得ている。

## 【0 1 2 8】

図 1 8、図 1 9 を用いて本実施例における領域識別結果の比較と修正の例を説明する。図 1 8 は二値化によって” COURSE” を含む文字行が全く抽出されなかった例であり、このような場合に図 1 9 に示した手順で領域識別を修正する。

## 【0 1 2 9】

まず、ステップ S 7 0 1 にてカラー領域識別結果 C を、ステップ S 7 0 2 にて二値領域識別結果 B を入力する。続いて、ステップ S 7 0 3 において行単位で双方の結果の対応を調べる。

【 0 1 3 0 】

次に、ステップ S 7 0 4 において片方には存在しない行があるかどうかをチェックする。このために、文字矩形の座標を比較して判断する。例えば、カラー領域識別から得られたある文字矩形に対して、二値領域識別結果にはこれと矩形座標が重なる座標が全くない場合に、片方には存在しない行があると判断する。

【 0 1 3 1 】

ステップ S 7 0 4 における判断結果が Y e s ならば、ステップ S 7 0 5 の判断へ進み、N o ならば、本処理を終了して図 1 4 のステップ S 6 0 7 の O C R 処理へ進む。

【 0 1 3 2 】

ステップ S 7 0 5 において、前記検出された文字行が二値領域識別結果 B にのみ存在し、カラー領域識別結果 C には該当する文字行は存在しないかどうかをチェックする。N o ならば、ステップ S 7 0 8 へ進む。

【 0 1 3 3 】

Y e s ならば、該行内の文字色のばらつきを調べるために、ステップ S 7 0 6 において原画像に立ち戻って文字色を特定し、ステップ S 7 0 7 において各文字の文字色に大きなばらつきがあるかどうかを判定する。ここでは、例えば各文字色（R G B 値）の分散を求めて、これが一定値以上大きい場合には色のばらつきが大きいと判断する。

【 0 1 3 4 】

ステップ S 7 0 5 の判断において Y e s である場合のようにカラー領域識別では抽出できなかった文字行が二値領域識別結果 C に存在する場合には、二値化が不十分であったために文字行にノイズが混入していることが多い。本発明では、そのような場合に、そのノイズの色と正しい文字色とは異なると仮定し、ステップ S 7 0 6、S 7 0 7 の処理を行うことによって色のばらつき具合を見ることでこの誤りを検出している。

## 【 0 1 3 5 】

ステップ S 7 0 7 の判断が N o、もしくはステップ S 7 0 5 の判断が N o ならば、ステップ S 7 0 8 において、該当する行を包含する領域を作成する。

## 【 0 1 3 6 】

ステップ S 7 0 7 の判断が Y e s ならば、図 1 8 に示す処理は終了して図 1 2 のステップ S 6 0 7 の O C R 処理に進む。

## 【 0 1 3 7 】

ステップ S 7 0 8 の処理の後、ステップ S 7 0 9 として該領域を再度二値化して、該領域を二値領域識別し、ステップ S 7 0 2 からの処理を再度行う。

## 【 0 1 3 8 】

上記の処理において、ステップ S 7 0 5 において N o である場合とは、カラー領域識別では抽出された文字行が二値領域識別の結果では存在しない場合であり、その場合、ステップ S 7 0 8 以降の処理で示すように、該当する文字行を包含するような領域を作成して、二値化および二値領域識別して、両者の領域識別結果に著しい差がなくなるまでこのフィードバック処理を繰り返して最終的な領域識別結果の向上を図ることとしている。

## 【 0 1 3 9 】

すなわち、カラー領域識別では文字行が抽出されたのにも関わらず、二値領域識別では対応する文字行が存在しなかった場合、文字色を調べることなくフィードバック処理を行って該文字行の再抽出を試みる処理を行うようにしている。

## 【 0 1 4 0 】

ここで、ステップ S 7 0 5 で N o の場合に文字色を厳密に調べることなくフィードバック処理に踏み切る理由は、色情報を直接用いたカラー領域識別のほうが、形状や二値の情報しか用いていない二値領域識別よりも文字列の有無を検出する精度が高いからである。この段階では、カラー領域識別でも正確に文字行を切り出すことはできない可能性があるが、前記理由から文字列が存在する可能性は高いので、二値化と二値領域識別を繰り返すことにより最終的な O C R のための文字抽出精度向上を図ることとしている。

## 【 0 1 4 1 】

図 2 0 に、領域識別結果の比較と修正の他の例を示す。図 2 0 ( a ) はカラー領域識別の結果例であり、本来の正しい行切り出し結果を示している。図 2 0 ( b ) は二値領域識別結果を示す図であり、二値化が不十分であったため、文字の一部が正しく抽出できずに周辺のノイズと融合して図・写真領域が誤抽出されている例を示している。

## 【 0 1 4 2 】

このような場合、本実施例では図 2 1 の手順で領域識別結果の正当性を検証して、二値化や領域識別処理を繰り返すフィードバック処理により誤りを修正する。

## 【 0 1 4 3 】

ステップ S 8 0 1 においてカラー領域識別結果を入力し、ステップ S 8 0 2 において二値領域識別結果を入力する。次に、ステップ S 8 0 3 として行単位で双方の結果の対応を調べる。

## 【 0 1 4 4 】

ステップ S 8 0 4 において、特徴の異なる行が所定の数以上連続して存在するかどうかをチェックする。本実施例では所定の行数を 3 行としている。図 2 1 の例では、 $L^c_i$  と  $L^b_i$  がそれぞれ対応する行になる。具体的には次の条件のいずれかを満たせばステップ S 8 0 4 における特徴が異なると判定される。

- ・ 行幅が 6 0 % 以下か？
- ・ 行の平均文字サイズの差が 3 p t 以上か？
- ・ 行の平均文字色（輝度の平均）の差が 3 0 以上か？

また、エッジの分布等を判定に用いてもよい。

## 【 0 1 4 5 】

ステップ S 8 0 4 にて Yes ならば、ステップ S 8 0 5 において該当する行を包含する領域を作成し、ステップ S 8 0 6 において該領域を再度二値化し、ステップ S 8 0 7 にて該領域を再度二値領域識別する。

## 【 0 1 4 6 】

ステップ S 8 0 4 にて No ならば、この処理は終了して図 1 4 のステップ S 6 0 7 の OCR 処理を行う。



## 【 0 1 4 7 】

OCR処理については、この時点でOCRに適した二値画像と領域識別結果が生成されているので、従来の方法でOCR処理を行う。OCR結果も従来と同様に文字コード、認識確信度、位置やフォントサイズなどのレイアウト情報と共にOCR結果DBに保存する。

## 【 0 1 4 8 】

なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

## 【 0 1 4 9 】

## 【発明の効果】

上記の通り、本発明によれば、従来の方法と異なりカラー画像から色情報を用いて、2値化することなく直接文字／罫線／イラスト及び写真等の領域を識別するので、色情報の欠落がなく、精度良く領域識別が可能になる。

## 【 0 1 5 0 】

また、本発明によれば、カラー文書画像の背景色を精度良く特定するので、カラー文書画像の背景色と異なる画素を統合してランの生成が可能になる。すなわち、文書要素として意味のある矩形を精度良く抽出することが可能となる。また、カラー文書画像の一部の部分領域に限定して背景色特定の処理を施すことにより、該部分領域内の再帰的な領域識別も可能になる。

## 【 0 1 5 1 】

また、効率良く原画像を縮小することで、画像へのアクセスコストを大幅に削減できる。すなわち、処理速度が向上し、使用メモリ量を削減することが可能となる。また、本発明における圧縮方法によれば、画像を高速に平滑化する効果もあり、文書要素の抽出が容易になる。また、縮小画像を作成する際に、画像をm×mの格子状にすることで、本発明の実装が容易になる。

## 【 0 1 5 2 】

また、背景領域と文書要素を分離する際には各色値の差異を用いるため、簡単な演算で画素間の相違度を求めることができ、十分な精度が得られる。

## 【 0 1 5 3 】